MICHEL QUEYSANNE

Maître-assistant de la Facultad
de Ciencias de París

Algebra Básica

Traducción de

JOSE LUIS VIVIENTE

Catedrático de la Facultad
de Ciencias de Zaragoza

Primera edición

editorial vicens—vives

PREFACIO

CONTENIDO Y METODO DEL LIBRO

El presente libro, dedicado al *Algebra*, está destinado, en primer lugar, a los estudiantes que preparan la licenciatura de *Matemáticas generales* y *Física* y el ingreso en las escuelas técnicas superiores en *Matemáticas especiales*. Igualmente se destina a toda persona que precise buenos conocimientos de base en Algebra en el ejercicio de su actividad profesional (físicos, ingenieros, técnicos). Por ello no se ha considerado un deber el limitarnos estrictamente a los programas oficiales de Matemáticas especiales. Este volumen comprende, pues:

- *El desarrollo de esos programas*, que son muy semejantes.

- Ciertas nociones que han sido estudiadas en los últimos cursos del bachillerato y no forman parte, *stricto sensu,* del programa universitario: *enteros naturales* (me he aprovechado de ello para demostrar, a partir de un pequeño número de axiomas, las propiedades de los *conjuntos finitos* a los que constantemente se hace alusión de modo consciente o no), *enteros racionales* y divisibilidad; igualmente he tratado de dar una definición correcta de los *ángulos* y de su medida.

- Un pequeño número de cuestiones que el lector podrá necesitar en sus estudios posteriores y que raramente son tratadas en otros textos, por la equívoca razón de que son cuestiones bien conocidas, y que en otro tiempo se encontraban en los programas usuales universitarios y de escuelas técnicas superiores (*grupo circular, polinomios y fracciones racionales simétricas, resultante*).

Debido a este contenido resulta que el libro es grueso. Naturalmente que ni el autor ni, creo yo, ningún profesor puede o querrá enseñar todas estas materias en un año. Pero, por el contrario, puedo decir que en el transcurso de mis quince años de enseñanza en escuelas técnicas superiores y de siete años en Matemáticas generales no existe una sola cuestión de las tratadas en

este libro que yo no haya expuesto por lo menos una vez a estudiantes del nivel considerado. Desde luego, siempre es delicado la fijación del límite entre lo que se debe tratar y lo que es preciso omitir.

Nos ha parecido conveniente en este libro de iniciación al *Algebra lineal* el estudio de los espacios vectoriales al constituir un todo coherente y tener aplicaciones suficientemente ricas y dejar de lado los módulos cuyo estudio puede desorientar a los principiantes. *A fortiori* no hemos hablado ni de producto exterior ni de producto tensorial, a pesar de que estas nociones sean, a los ojos de quien las conoce, subyacentes a las de determinante y de forma bilineal.

Por el contrario, en el estudio de los *grupos, anillos, cuerpos, espacios vectoriales,* hemos hecho una amplia llamada a las nociones de *relación de equivalecnia compatible con la estructura estudiada,* de *homomorfismo,* de *subgrupo distinguido* y de *ideal,* porque aclaran profundamente cuestiones que, indiscutiblemente, están en el centro del programa.

Tampoco hemos querido limitarnos a la definición de los *ideales* de **Z** y de K[X], porque no siendo principales una gran cantidad de anillos, como K[X, Y], e incluso no factoriales, como $\mathbf{Z}[i\sqrt{5}]$, nos ha parecido que sería falsear el espíritu de los estudiantes dejarles creer que sólo existen ideales principales.

A propósito de dos nociones principales del libro: el anillo de polinomios K[X] y el anillo $\mathcal{L}(E)$ de los endomorfismos de un espacio vectorial, hemos presentado la noción de *álgebra* sobre un cuerpo conmutativo.

En fin, dada la importancia de las *formas bilineales* y *hermitianas* en Matemáticas y en Física su estudio ha sido tratado con bastantes detalles (en dimensión finita).

Para limitar las dimensiones del presente volumen, el estudio de los *espacios afines* ha sido reservado como introducción a lo que tradicionalmente se llama la *Geometría analítica.*

El método empleado en la exposición de estas nociones tiene en cuenta la diferencia entre un *libro de texto* y un *tratado*: en un libro de texto la buena pedagogía exige no introducir las nociones estudiadas hasta el momento de su utilización. De ello resulta una gran dispersión de las definiciones, dispersión molesta para el estudiante que revisa o repasa, y sobre todo para el lector que consulta el libro. Por el contrario, en una exposición sistemática se tiene la tendencia de agrupar definiciones y propiedades relativas a una misma noción, sin que el interés de ciertas definiciones pueda ponerse de manifiesto hasta bien desarrollado el libro. En esta obra hemos adoptado una solución intermedia: exponer de modo agrupado las nociones fundamentales (capítulos 1, 2 y 3), sin preocuparnos de su utilización inmediata; pero cada vez que las hemos utilizado posteriormente enviamos al lector explícitamente a su definición, y con frecuencia hemos repetido y comentado ésta: en ello se encuentra una segunda razón para las dimensiones del libro.

El capítulo primero debe ser previamente estudiado de modo detenido y profundo: no se puede hacer Matemáticas sin conocerlo bien; por otra parte, la modernización de los programas de la enseñanza de segundo grado

la enseñanza media lo hace menos nuevo de lo que hubiera sido, hace algunos años, para los estudiantes del propedéutico o selectivo.

El capítulo 2 contiene, aparte del Análisis combinatorio, una teoría de los números enteros naturales y de los conjuntos finitos, cuyos resultados son bien conocidos y que puede ser saltada en primera lectura.

En cuanto al capítulo 3 (Leyes de composición) para un principiante puede ser interesante no leer más que la primera sección, sin que ello sea inconveniente para más adelante *volver* de nuevo a este capítulo para encontrar una síntesis que permita comprender mejor la unidad de los capítulos siguientes.

Realizado el estudio de las estructuras fundamentales (grupos, anillos, cuerpos, espacios vectoriales) y números complejos (capítulos 4 al 7), queda por desarrollar, por una parte, el Algebra lineal, los determinantes y sus aplicaciones (capítulos 7 al 10) y, por otra parte, el estudio de los polinomios, de las fracciones racionales y el de las ecuaciones algebraicas (capítulos 11 al 14). Lógicamente, el lector puede intercambiar el orden de estos dos últimos grupos de capítulos, *pero entonces él se verá privado de toda la ayuda que puede proporcionarle el Algebra lineal al estudio de los polinomios.*

La reducción de los endomorfismos de un espacio vectorial de dimensión finita y el estudio de las formas bilineales y hermitianas, al precisar el estudio previo de los numerosos capítulos anteriores, terminan el libro (capítulos 14 y 15).

En fin, sin temor a alargar aún más el libro nos hemos entretenido con numerosas observaciones sobre el por qué de la elección de las definiciones y sobre la significación de los resultados enunciados, ilustrándolos con numerosos ejemplos y ejercicios. Estos últimos son de dos clases: los unos, situados en cada parágrafo, son en su casi totalidad muy fáciles, ilustran una cuestión precisa o bien son aplicaciones inmediatas de la misma; algunos, aunque muy pocos, prolongan, por el contrario, los resultados del texto; unos y otros se desarrollan bajo el título *Ejemplos y ejercicios.* Normalmente el lector debería estudiarlos al mismo tiempo que el texto propiamente dicho. Al final de cada capítulo se proponen otros ejercicios y problemas más largos que versan sobre la materia del capítulo o de las materias de capítulos anteriores; los más difíciles *están señalados con un* asterisco(*), aunque esta calificación de ejercicio difícil es desde luego muy subjetiva.

En los ejemplos y ejercicios situados en los párrafos, no he tenido inconveniente alguno en utilizar extensamente los conocimientos de Matemáticas elementales, incluso si las nociones correspondientes no se hallaban definidas hasta más adelante en este libro (enteros, racionales, reales, complejos) o incluso no se tratan en él (funciones continuas, funciones derivables, integral definida, vectores libres, desplazamientos). No existe, en efecto, inconveniente *lógico* alguno en ilustrar las nociones de relación de equivalencia y de conjunto cociente mediante la noción de entero módulo p y a dar en seguida la definición de los enteros racionales con la ayuda de una relación de equivalencia. Como hemos hecho observar más arriba a propósito del capítulo 3, *el llegar a comprender totalmente un libro de Matemáticas exige frecuentes vueltas a lo anterior:* una definición no puede ser completamente comprendida más que

si no dispone de ejemplos a los que ella precede con frecuencia en una exposición lógica única.

Hemos dedicado un gran cuidado a la elección de los *términos* y *notaciones*. De un modo general hemos adoptado las denominaciones y las notaciones de N. BOURBAKI, al menos tal como se las puede conocer en los últimos fascículos aparecidos. Y esto por dos razones: por una parte, porque son cada vez más utilizadas en todo el mundo, e inmediatamente después porque son el fruto de serias discusiones colectivas, que garantizan, frecuentemente, su carácter más correcto. Naturalmente el tratado de N. BOUBARKI no es directamente accesible a los estudiantes de propedéutico; pero debo decir que en veinte años de enseñanza, cuando yo he encontrado una dificultad en la elección de una "buena noción" o en la elección de una definición o de un término, es casi siempre en N. BOURBAKI donde he encontrado la solución. Pero igualmente hemos dado denominaciones y notaciones que el lector podrá encontrar en otros textos o tratados.

Además de los fascículos de N. BOURBAKI, he utilizado mucho los libros y cursos policopiados de Algebra aparecidos estos últimos años, en particular los de MM. CHEVALLEY, CHOQUET, DIXMIER, DUBREIL, GODEMENT, PISOT, LICHNEROWICZ y ZAMANSKY, sin olvidar el libro de VAN DER WAARDEN, que hoy es aún, bajo un volumen reducido, el tratado de Algebra más completo.

Para los ejercicios, he utilizado mucho, tanto obras antiguas como recientes, franceses y extranjeras. También he tomado algunas de las colecciones de problemas dados estos últimos años en la Facultad de París; mi agradecimiento a todos mis compañeros profesores-adjuntos y adjuntos por la colaboración que así me han prestado.

Agradezco vivamente a mi antiguo camarada y amigo ANDRÉ REVUZ los constantes ánimos y preciosos consejos que me ha prodigado, así como a madame REVUZ, que redactó el curso de A. P. M., desarrollado por su marido, cuyo texto y ejercicios me han sido muy útiles.

A mi colega mademoiselle BRETIN quiero expresar también mi agradecimiento por su ayuda en el ingrato trabajo de la corrección de las pruebas.

M. Q.

Nota.—Cada capítulo está dividido en general en secciones (I, II, ...). Las secciones están divididas en párrafos numerados de principio a fin del libro del 1 al 240. Las referencias envían a los párrafos (ejemplo: ver § 25). Los ejemplos y ejercicios incluidos en el texto están numerados por párrafos (ejemplo: § 137, ej. 4). Los ejercicios de fin de capítulo están numerados consecutivamente de un extremo a otro del libro (ejemplo: ej. 208, al final del capítulo 8).

INDICE

Página

1

CONJUNTOS APLICACIONES. RELACIONES

I. Introducción. Nociones de lógica. Conjunto (elemento, pertenencia).
II. Inclusión. Reunión. Intersección.
III. Producto cartesiano. Relaciones. Correspondencias.
IV. Aplicaciones de A en B.
V. Relaciones de equivalencia.
VI. Relaciones de orden.
VII. Conclusión.

I. Introducción. Nociones de pertenencia. Conjunto (elemento, pertenencia)

1. Introducción

Una teoría matemática se presenta bajo la forma de una sucesión de enunciados (definiciones y proposiciones) tales que toda definición viene dada mediante términos ya definidos y toda proposición demostrada mediante proposiciones ya admitidas. Es este orden lógico en los enunciados el que transforma una colección de resultados, sin relación alguna entre ellos, en una *teoría deductiva*: el primer ensayo de constitución de una tal teoría se debe a EUCLIDES, o a sus predecesores inmediatos (siglos IV y III antes de J.C.).

Estas consideraciones exigen varias aclaraciones:

Las definiciones, las proposiciones y sus demostraciones son enunciadas con las *palabras de una lengua* (el griego para EUCLIDES, el español para nosotros) dejándoles su sentido ordinario, si ninguna confusión puede originarse, precisando ciertos términos en el caso contrario.

Las demostraciones se efectúan mediante las *reglas de la lógica*. La historia del pensamiento muestra, por otra parte, que la elaboración de la lógica

y la constitución de los primeros "Elementos" de Matemáticas se han efectuado conjuntamente y que en cada momento de la historia en que una preocupación de mayor rigor se ha manifestado (siglo XVII, fin del XIX y principio del XX) hubo una acción recíproca entre la lógica y la matemática (los matemáticos piensan, generalmente, que es el progreso de la matemática lo que ha favorecido el progreso de la lógica, pero puede que sea un punto de vista subjetivo...).

Señalaremos, no obstante, que la imagen dada anteriormente de una teoría matemática es en parte falsa, ya que inicialmente se presentan los *términos primitivos* que no se definen y las proposiciones primitivas (*axiomas*) que no se demuestran; términos y proposiciones que constituyen lo que M. DANIEL LACOMBE ha llamado *"dominio intuitivo de base"* de la teoría considerada[1].

La preocupación siempre creciente del rigor lleva consigo el retroceso de este dominio en el desarrollo histórico de las Matemáticas. Asimismo en una misma época, la nuestra, por ejemplo, este dominio intuitivo de base en una teoría expuesta a un auditorio (alumnos, estudiantes, investigadores) disminuye con la madurez intelectual del auditorio.

Una de las mayores preocupaciones de los matemáticos ha sido reducir lo más posible este "residuo intuitivo"; salvando todo lo que puede tener de excesiva limitación el esquema siguiente, se pueden distinguir tres etapas en la historia de las Matemáticas:

—*Matemáticas clásicas* (hasta el siglo XIX, último tercio inclusive). Todos los términos primitivos y axiomas no están siempre expuestos de un modo claro; se recurre ampliamente de manera explícita o implícita —lo que es más grave— a la intuición para las demostraciones. El recurso a "nociones comunes" que constituyeron más tarde la teoría de los conjuntos es muy utilizada con la lógica clásica (la de ARISTÓTELES y los Escolásticos).

—*Matemáticas axiomatizadas* (desde el final del siglo XIX). Bajo la influencia de PÉANO y sobre todo ds HILBERT se subsanan en parte las insuficiencias señaladas más arriba. Los términos primitivos y los axiomas de cada teoría son enunciados de un modo preciso. Pero el empleo de la lengua corriente y de las reglas de una lógica perfeccionada hacen aún uso implícito de la intuición; de donde la existencia de un residuo intuitivo —relativo en particular a la teoría de los conjuntos, base de toda matemática—, bastante considerable para suscitar numerosas paradojas que alimentan las controversias entre matemáticos. Estas paradojas serán poco a poco eliminadas al precisar de modo conveniente la noción de conjunto (ver § 3).

—*Matemáticas formalizadas* (primera mitad del siglo XX). Entonces fue tan grande la tentación de sustituir los objetos matemáticos por puros símbolos sin contenido intuitivo y los enunciados por simples reglas de empleo de estos símbolos: de este modo se encuentran fundidas a la vez e inexplicable-

(1) D. LACOMBE, *Les idées actuelles sur la structure des mathématiques.* XXe Semaine de Synthèse. París, 1956.

mente mezcladas la lógica y la teoría de los conjuntos, habiendo sido excluido de la matemática todo juicio sobre estas teorías y relegado a la *metamatemática.* Esta última ciencia puede ser inmediatamente formalizada y así sucesivamente: a cada etapa, el contenido del "dominio intuitivo de base" se reduce; no nos corresponde afirmar si esta concepción formalista ha conseguido reducirlo a la nada.

La actitud adoptada en este texto será definida precisando nuestro "dominio intuitivo de base":

— *reglas de la lógica* de las proposiciones convenientemente precisadas;

— nociones de *conjunto,* de *elementos,* de *pertenencia,* intuitivamente admitidas, ilustrándolas con ejemplos y precisándolas, como se dirá en el § 3;

— admitir la existencia del *conjunto de los enteros naturales* **N** (esto para simplificar el curso, pues esta existencia resulta de la teoría de conjuntos).

Presentados estos preliminares en esta sección I del capítulo 1 y en la sección I del capítulo 2, el resto del curso se desarrolla con todo el rigor compatible con la madurez de nuestros estudiantes: pues no se trata en nuestro nivel, después de todo muy elemental respecto al conjunto de las Matemáticas, de eliminar completamente la *intuición* de nuestros estudiantes(*) como ellos la entienden, sino de recurrir a ella únicamente en *casos explícitos.*

Naturalmente utilizaremos nuestra lengua maternal convenientemente precisada, separando claramente *el texto, que debe estar redactado en español correcto, y el lenguaje simbólico (fórmulas) relativo a cada teoría.* Cuando no pueda surgir ninguna ambigüedad emplearemos expresiones o notaciones simplificadas (*abuso de lenguaje, abuso de notaciones*) destinadas a simplificar el texto; cada vez que así lo hagamos procuraremos indicarlo.

1. Nociones de lógica

Una *aserción* es un enunciado del que se puede afirmar sin ambigüedad si es cierto o si es falso.

Por ejemplo, «$3 < 10$» es una aserción verdadera; «$5 < 2$» es una aserción falsa. «Todo triángulo isósceles es equiángulo» es una aserción verdadera.

Los enunciados que encontraremos a menudo son más generales: el enunciado será *verdadero en ciertos casos, falso en otros,* pero para una *situación dada* podremos decidir *si un enunciado es verdadero o si es falso.* Este enunciado se llama una *proposición.* Representaremos una proposición por una letra P, Q, R ...

Esta actitud elemental no prejuzga la existencia o la no existencia de enunciados de los que no se sabría demostrar si son verdaderos o falsos.

Se ve que una aserción es una proposición siempre verdadera o siempre falsa. En lugar de escribir "P es verdadera" escribiremos solamente "P": es

(*) *N. del T.* — Tal como el estudiante entiende tal palabra.

decir, *cuando escribimos "P" nos colocamos en una situación en que "P" es cierta.*

Por ejemplo, «$x < 10$» es una proposición, es verdadera para los números estrictamente inferiores a 10, falsa en los demás casos.

«La altura del triángulo T es mediana del triángulo T» es una proposición verdadera para los triángulos T isósceles, falsa en los demás casos.

La *negación* de una proposición "P" que escribiremos "*no* P" es verdadera cuando P es falsa, falsa cuando P es verdadera[2].

La negación de una proposición se puede esquematizar en la tabla I, que se llama una *tabla de verdad*:

P	no P
V	F
F	V

TABLA I

P 1	Q 2	P y Q 3	P o Q 4	$P \Rightarrow Q$ 5	$P \Leftrightarrow Q$ 6
V	V	V	V	V	V
V	F	F	V	F	F
F	V	F	V	V	F
F	F	F	F	V	V

TABLA II

La tabla I se lee fácilmente: V significa verdadero y F falso. Consideremos ahora dos propiedades P, Q. Las situaciones respectivas en que son verdaderas o falsas conducen a cuatro situaciones posibles indicadas por cada línea de las dos primeras columnas de la tabla II. Podemos definir una nueva proposición R indicando en otra columna el valor de R (V o F) correspondiente a cada una de estas cuatro situaciones: se pueden definir así dieciséis proposiciones nuevas (ver ejercicio 1 más abajo). En las columnas 3 a la 6 de la tabla II hemos indicado las tablas verdaderas de las cuatro proposiciones más usadas deducidas de las proposiciones P, Q.

La *conjunción* de dos proposiciones P, Q (col. 3) que escribiremos "P y Q" es verdadera si y solamente si P y Q son simultáneamente verdaderas y en todos los demás casos falsas. Dos proposiciones son *incompatibles* si su conjunción es siempre falsa. Por ejemplo, las proposiciones P, no P, son incompatibles, pero éste es un caso particularísimo de incompatibilidad.

El punto de vista intuitivo consiste en considerar como siempre falsa la proposición "P y (no P)" ("*principio de no contradicción*").

Así $x < 3$ y $x > 5$ son incompatibles.

La *disjunción* de dos proposiciones P, Q (col. 4) que escribiremos "P *o* Q" es verdadera si al menos una de las proposiciones P, Q es verdadera, falsa en

(2) Se designa también la negación de P: $\sim$ P, $\bar{P}$, P̸.

los demás casos (es decir, si y sólo si P y Q son falsas simultáneamente). Se observa que el sentido de la palabra "*o*" precisa el del lenguaje corriente, para el que "*o*" tiene dos sentidos:

— el que hemos dado;

— el sentido exclusivo: *o bien* P es verdadero (y Q es falsa) *o bien* Q es verdadera (y P es falsa).

En Matemáticas tomaremos siempre el sentido indicado más arriba: veremos más adelante una de las razones de esta elección.

El punto de vista intuitivo consiste en considerar siempre verdadera la proposición "o bien P, o bien no P" ("*principio del tercio excluso*").

La proposición "(*no* P) *o* Q" llamada *implicación* (col. 5) se designa

$$P \Rightarrow Q \tag{1}$$

y se enuncia: "P *implica* Q" o "P *entraña* Q". Siendo P y Q aserciones si "$P \Rightarrow Q$" es verdadera, se dice que es un *teorema* (es decir, una aserción demostrada en la que P es la *hipótesis* y Q la *conclusión*.

Si P y Q dependen de variables y si "$P \Rightarrow Q$" es verdadera, no importa cuáles sean los valores atribuidos a las variables, se dice indistintamente:

P es una condición *suficiente* de Q
Q es una condición *necesaria* de P

Observemos que las dos últimas líneas de la tabla de verdad de "$P \Rightarrow Q$" muestran que el sentido que hemos dado a las palabras "implica" y "entraña" es más general que el del lenguaje corriente:

Así «4 es un número primo» $\Rightarrow$ «Madrid es la capital de España».
«Barcelona es la capital de España» $\Rightarrow$ «2 es un número primo» con dos implicaciones (verdaderas).

Dos proposiciones son *equivalentes* si cada una de ellas implica la otra (col. 6): las situaciones en que P es verdadera [resp. falsa] son exactamente las mismas situaciones en que Q es verdadera [resp. falsa]; se escribe

$$(P \Rightarrow Q) \quad y \quad (Q \Rightarrow P)$$
$$P \Leftrightarrow Q \tag{2}$$

en este caso los dos teoremas: "$P \Rightarrow Q$" y "$Q \Rightarrow P$" son verdaderos simultáneamente, se dice que son recíprocos uno del otro. Por ejemplo, en las dos equivalencias siguientes

$$no\ (no\ P) \Leftrightarrow P$$
$$(P \Rightarrow Q) \Leftrightarrow (no\ Q \Rightarrow no\ P)$$

la implicación del segundo miembro se llama la *implicación contrapuesta* de la

del primer miembro. Para dos proposiciones equivalentes P, Q diremos indiferentemente:

Para que Q [resp. P] sea verdadera *es necesario y suficiente* que P [resp. Q] sea verdadera.

Q [resp. P] es verdadera *si y sólo si* P [resp. Q] es verdadera.

La verdad de Q [resp. P] es *una condición necesaria y suficiente* de la verdad de P [resp. Q].

La implicación (1) es *transitiva*, es decir,

$$[(P \Rightarrow Q) \quad \text{y} \quad (Q \Rightarrow R)] \Rightarrow (P \Rightarrow R)$$

y análogamente para la equivalencia (2)

$$[(P \Leftrightarrow Q) \quad \text{y} \quad (Q \Leftrightarrow R)] \Rightarrow (P \Leftrightarrow R).$$

Señalemos finalmente que

$$(P \Rightarrow Q) \quad \text{y} \quad (Q \Rightarrow R) \quad \text{se designa} \quad P \Rightarrow Q \Rightarrow R$$
$$(P \Leftrightarrow Q) \quad \text{y} \quad (Q \Leftrightarrow R) \quad \text{se designa} \quad P \Leftrightarrow Q \Leftrightarrow R$$

Por otra parte, dadas dos proposiciones P, Q, las dos equivalencias siguientes, conocidas por el nombre de *reglas de dualidad*, son verdaderas

$$\textit{no}\ (P\ \textit{y}\ Q) \Leftrightarrow (\textit{no}\ P)\ \textit{o}\ (\textit{no}\ Q)$$
$$\textit{no}\ (P\ \textit{o}\ Q) \Leftrightarrow (\textit{no}\ P)\ \textit{y}\ (\textit{no}\ Q)$$

la simetría presentada por esas dos reglas es una de las razones de la elección del sentido de la palabra "*o*" en Matemáticas.

Razonamiento por reducción al absurdo. — Supongamos que se quiere probar que la proposición P es verdadera; el razonamiento consiste en introducir la proposición *no* P que se designará por P′ después a demostrar una implicación tal como

$$P' \Rightarrow Q'$$

donde Q′ es la negación de una proposición Q, *de la que se sabe que es verdadera.* Ahora bien,

$$(P' \Rightarrow Q') \Leftrightarrow (\textit{no}\ Q' \Rightarrow \textit{no}\ P') \Leftrightarrow (Q \Rightarrow P)$$

siendo Q *verdadera, resulta que* P *es verdadera.*

Recordemos que se presenta a menudo este razonamiento diciendo: supongamos que P sea falsa, es decir, P′ verdadera, de la implicación $P' \Rightarrow Q'$ resultaría que Q′ es verdadera; ahora bien, se sabe que Q es verdadera y el principio de la no contradicción permite entonces afirmar que Q′ es falsa. Según esto una proposición falsa no puede ser implicada por una proposición verdadera, luego P′ es falsa, es decir, P es verdadera (*principio del tercio excluso*).

La introducción de la hipótesis «P es falsa» *en realidad no ha servido para nada,* ya que la demostración por reducción al absurdo consiste en establecer la implicación $P' \Rightarrow Q'$ que no depende de la verdad o no verdad de P.

EJERCICIOS

1. Construir la tabla de verdad de las 16 proposiciones que se pueden obtener a partir de dos proposiciones P y Q. Enunciar en lenguaje corriente el sentido de cada una.

2. Verificar que la conjunción y la disjunción son asociativas, es decir, que

$$\text{P y (Q y R)} \Leftrightarrow \text{(P y Q) y R}$$
$$\text{P o (Q o R)} \Leftrightarrow \text{(P o Q) o R.}$$

3. Comprobar que la conjunción y la disjunción son distributivas una respecto la otra, es decir, que

$$\text{P y (Q o R)} \Leftrightarrow \text{(P y Q) o (P y R)}$$
$$\text{P o (Q y R)} \Leftrightarrow \text{(P o Q) y (P o R).}$$

4. Determinar las proposiciones equivalentes a

$$\text{(P o Q) y (R o S)}$$
$$\text{(P y Q) o (R y S).}$$

5. Siendo x e y números reales, resolver el sistema

$$\begin{cases} (x-1)(y-2) = 0 \\ (x-2)(y-3) = 0. \end{cases}$$

1. Noción de conjunto

a) Tanto los seres físicos (piedra, alumno, perro, mesa ...) como los objetos de nuestro pensamiento (número, función ...) que representaremos por las letras a, b, ..., α, β, ..., X, ... se considerarán bien definidos si poseemos un criterio que permita afirmar que dos de estos objetos representados por a y por b son o bien *idénticos* o bien *distintos*; escribiremos

$$(1) \quad a = b \qquad (2) \quad a \neq b$$

en el primer caso diremos que a y b son *iguales*, en el segundo que a y b son desiguales. Las dos proposiciones (1) (*igualdad*) y (2) (*desigualdad*) son cada una la negación de la otra. En el final del capítulo (§ 26, *c*) precisaremos esta noción de igualdad y el abuso de lenguaje de que es objeto. Basta decir que la igualdad $a = b$ significa que habiendo considerado al objeto de dos maneras distintas ha podido recibir dos "denominaciones" distintas representadas por a y b; reconociendo seguidamente que se trataba del mismo objeto se escribe $a = b$.

Por ejemplo, se calcula el número $5 + 7$ se le llama al resultado a, se calcula el número 3×4 se le llama b, las reglas de aritmética permiten decir $a = b$.

La noción de *conjunto* corresponde a las nociones corrientes de *conjunto*, de *colección*, de *agrupación*, de *clase*, etc., de objetos de cualquier naturaleza (como los considerados al principio del párrafo); en el lenguaje corriente estos objetos se llaman *elementos*, *miembros*, *individuos* ... del conjunto, de la colección, de la agrupación ...

Los matemáticos han escogido las palabras *conjunto* y *elemento* y dicen:

Un conjunto está constituido de elementos, palabras ambas que están precisadas por las siguientes reglas:

1. *Un conjunto* E *está bien definido cuando se posee un criterio que permite afirmar si el objeto a pertenece al conjunto* E *o no pertenece a dicho conjunto; se escribe y se lee, respectivamente,*

(3) $a \in E$	(4) $a \notin E$
"a pertenece a E"	"a no pertenece a E"
o "a es un elemento de E"	"a no es un elemento de E"
o "E contiene a"	"E no contiene a"

La fórmula (3) interpreta la proposición llamada *pertenencia* de un elemento a un conjunto, (4) su negación.

Las colecciones, agrupaciones, etc., considerados en el lenguaje corriente no verifican siempre este criterio; por ejemplo, la clase de las "personas rubias". En las clasificaciones que hace la Historia Natural: reino, rama, clase, encuentra seres difíciles de situar en una de estas agrupaciones, la pertenencia a una de ellas resulta ambigua: estas clasificaciones no son conjuntos en el sentido matemático.

2. *Un mismo ser matemático no puede ser a la vez un conjunto y un elemento de este conjunto, es decir, no nos permitimos escribir* $a \in a$.

3. *La colección de todos los conjuntos imaginables no es un conjunto,* si se llegase a tratar este caso, cosa poco frecuente (ver § 20), diremos la "*clase de todos los conjuntos*". No es posible aplicarle las propiedades de los conjuntos que vamos a demostrar y las operaciones que vamos a definir sobre los conjuntos.

Es el desconocimiento de estas reglas lo que ha motivado las paradojas señaladas en el § 1.

b) *Dos conjuntos son idénticos* (*o iguales*) si están constituidos de los mismos elementos; si no se les dice *distintos* (o *desiguales*), y se escribe

$$(5) \quad E = F \qquad (6) \quad E \neq F.$$

Aparte de estas reglas, acordamos una gran libertad en la definición de los conjuntos, como lo demuestran los siguientes ejemplos:

c) EJEMPLOS

1. El conjunto **N** de los *enteros naturales* 0, 1, 2 ...

El conjunto **Z** de los *enteros racionales* ... —2, —1, 0, 1, 2 ...

El conjunto **Q** de los *números racionales.*

El conjunto de todos los «números» utilizados en matemáticas elementales: enteros, racionales, $\sqrt{2}$, π ... que se llama conjunto de los *números reales,* representado por **R**.

Tenemos

$$3 \in \mathbf{N} \qquad 3 \in \mathbf{Z} \qquad 3 \in \mathbf{Q} \qquad 3 \in \mathbf{R}$$

$$-2 \in \mathbf{Z} \qquad \frac{5}{2} \in \mathbf{Q} \qquad \pi \in \mathbf{R},$$

pero

$$-2 \notin \mathbf{N} \qquad \frac{5}{2} \notin \mathbf{Z} \qquad \pi \notin \mathbf{Q}$$

luego estos conjuntos **N**, **Z**, **Q**, **R** son distintos

2. El *conjunto de todos los puntos de un plano.*

3. El conjunto de todos los triángulos isósceles es igual al conjunto de todos los triángulos que tienen dos ángulos iguales.

El conjunto de todos los triángulos isósceles es distinto del conjunto de los triángulos rectángulos.

4. El conjunto de estudiantes inscritos en el primer curso de Matemáticas el 1 de octubre de 1967 en la Facultad de Zaragoza.

5. El conjunto cuyos elementos son:

«Las tres Gracias, todos los enteros naturales pares, el perro perteneciente al autor de este libro el 1 de febrero de 1965, la masa volúmica del mercurio, los estudiantes que han obtenido matrícula de honor en el primer curso de *Matemáticas* en 1966.»

Se puede concebir que un signo, una letra, por ejemplo, designe un *elemento particular* de un conjunto E, o al contrario que designe un *elemento cualquiera* de un conjunto E; este último caso se enunciará indistintamente:

"sea x un *elemento cualquiera* de E";

"sea x un *elemento arbitrario* de E";

"sea x un *elemento genérico* de E";

se dice también que: "x *es una variable que describe* E", o simplemente: "x *describe* E".

d) Los conjuntos de la vida corriente, por grandes que sean, son siempre *finitos*: precisaremos esta palabra en el capítulo 2. En Matemáticas consideraremos los conjuntos no finitos llamados *infinitos,* por ejemplo, **N**, el conjunto de los puntos de un plano.

Un conjunto finito puede estar definido por la enumeración de sus elementos; por ejemplo, el conjunto definido más arriba en el número 4. Si E tiene por elementos a, b, c, d, e, se escribirá

$$E = \{a, b, c, d, e\}$$

el orden en que se escriben los elementos es indiferente, así

$$\{a, b\} = \{b, a\}.$$

Nos vemos así conducidos a considerar los conjuntos de un solo elemento a, lo escribiremos $\{a\}$ luego

$$a \in \{a\} \quad \text{y} \quad b \neq a \Leftrightarrow b \notin \{a\}.$$

Esta distinción entre a y $\{a\}$ es fundamental; si no se hiciera se violaría la regla 2) antes citada. Corresponde a una distinción hecha en la vida corriente: si en un instituto hay un solo alumno que estudia el *vasco,* hay que distinguir entre *este alumno* y la *clase* de *vasco* que no comprende más que este alumno.

Se puede también considerar como conjunto a un conjunto que no tenga ningún elemento; se le llama *conjunto vacío* y se representa por $\varnothing$; así cualquiera que sea a

$a \in \varnothing$ es siempre falsa

$a \notin \varnothing$ es siempre verdadera.

II. Inclusión. Reunión. Intersección

4. Inclusión. Parte. Complementario. Conjunto de las partes

a) Diremos que un conjunto F está *incluido* en un conjunto E cuando *todo elemento* de F *pertenece* a E; se escribe $E \subset F \equiv \forall x \in E \to x \in F$

$$(1) \quad F \subset E \qquad \text{o} \qquad (1') \quad E \supset F$$

por definición

$$(2) \quad (F \subset E) \Leftrightarrow (x \in F \Rightarrow x \in E).$$

La fórmula (1) se lee indistintamente:

"F *está incluido en* E";
"F *es una parte de* E";
"F *es un subconjunto de* E",

y la fórmula (1'):

"E *contiene* F";
"E *admite* F *como parte*";
"E *es un superconjunto de* F".

La expresión, con frecuencia empleada, "E *contiene* F" no se recomienda por la posible ambigüedad con la expresión "E *contiene a*", siendo a un elemento de E.

Las fórmulas (1) y (1') traducen de dos maneras distintas una misma proposición que se llama *inclusión* de conjuntos.

La fórmula (2) muestra inmediatamente que la inclusión es:

— *reflexiva,* es decir, $E \subset E$ para todo conjunto E;
— *antisimétrica,* es decir, $[(E \subset F) \text{ y } (F \subset E)] \Rightarrow E = F$;
— *transitiva,* es decir, $[(E \subset F) \text{ y } F \subset G] \Rightarrow E \subset G$.

La reflexividad de la inclusión muestra que la *igualdad de los conjuntos es un caso particular de la inclusión*; se emplea el término de *inclusión estricta*(3) para caracterizar el caso en que

$$F \subset E \quad y \quad E \neq F.$$

E es, pues, una parte de E, se le llama la *parte plena* de E. La definición que hemos dado de la implicación (ver § 2) muestra que, para todo x y todo conjunto E,

$$x \in \varnothing \Rightarrow x \in E,$$

luego para todo conjunto E: $\varnothing \subset E$, es decir, el conjunto vacío es una parte de todo conjunto, se le llama la *parte vacía* de E.

Una parte no vacía de E, distinta de E, se llama *parte propia* de E.

Los conjuntos considerados en el ejemplo 1 del § 3 verifican

$$\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R},$$

cada uno de ellos es una parte propia de los siguientes.

Cuando existe al menos un elemento de F que no pertenece a E, se dice que F no está incluido en E y se escribe

$$F \not\subset E.$$

ATENCIÓN: $F \not\subset E$ no es de ningún modo equivalente a $E \subset F$ (ver § 5, ejercicio 6).

b) Dada una parte A de E, se llama *complementario de* A *con respecto a* E el conjunto de elementos de E que no pertenecen a A; se le designa $\complement_E A$ (o $\complement A$ si no ha lugar a duda) o bien $E - A$.

Tenemos, cualquiera que sean un conjunto E y una parte A de E,

$$B = \complement_E A \Leftrightarrow A = \complement_E B,$$

dicho de otro modo

$$\complement_E \left(\complement_E A \right) = A,$$

en particular

$$\complement_E E = \varnothing \qquad \complement_E \varnothing = E.$$

(3) Ciertos autores emplean las notaciones $F \subseteq E$ para la inclusión y $F \subset E$ para la inclusión estricta. Nosotros no lo haremos, escogiendo el símbolo más simple para la inclusión, que se presenta con bastante más frecuencia que la inclusión estricta.

Los complementarios respectivos de $\{0\}$ con relación a los conjuntos **N**, **Z**, **Q**, **R** se designarán

$$\mathbf{N}^*,\ \mathbf{Z}^*,\ \mathbf{Q}^*,\ \mathbf{R}^*$$

que se lee: "**N** estrella, **Z** estrella, ...".

c) Consideremos todas las partes de un conjunto E; describen un nuevo conjunto llamado *conjunto de las partes* de E designado $\mathfrak{P}(E)$; se tiene, pues,

$$A \subset E \Leftrightarrow A \in \mathfrak{P}(E)$$

en particular si a es elemento de E (no vacío)

$$a \in E \Leftrightarrow \{a\} \subset E \Leftrightarrow \{a\} \in \mathfrak{P}(E).$$

Cualquiera que sea E tenemos

$$\varnothing \in \mathfrak{P}(E) \qquad E \in \mathfrak{P}(E),$$

aunque E sea vacío, $\mathfrak{P}(E)$ no lo es: contiene al elemento $\varnothing$ como elemento único.

EJERCICIOS

1. Determinar los elementos de $\mathfrak{P}(E)$ para $E = \{a, b, c, d, e\}$, a, b, c, d, e siendo distintos dos a dos.
2. Determinar $\mathfrak{P}(E)$ y $\mathfrak{P}(\mathfrak{P}(E))$ para un conjunto de dos elementos.
3. Determinar $\mathfrak{P}(E)$, $\mathfrak{P}(\mathfrak{P}(E))$ y $\mathfrak{P}(\mathfrak{P}(\mathfrak{P}(E)))$ para un conjunto de un elemento.
4. Teniendo E n elementos, ¿cuál es el número de los elementos de $\mathfrak{P}(E)$? (razonar por recurrencia).

Se pueden representar los conjuntos por *esquemas*; estos esquemas sirven solamente para ayudar a la intuición, no pueden de ningún modo utilizarse para demostrar enunciados sobre los conjuntos; así el esquema (1) es una "figura" representando la inclusión; el esquema (2) una "figura" representando la transitividad de la inclusión.

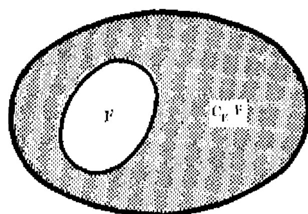


Fig. 1

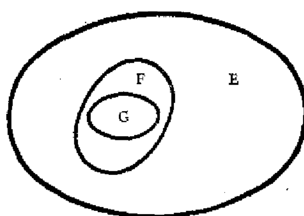


Fig. 2

5. Intersección y reunión de conjuntos

Se llama *intersección* de dos conjuntos E, F y se designa E ∩ F, leyéndose "E *intersección* F" *el conjunto descrito por los elementos comunes a* E *y a* F, luego

$$x \in E \cap F \Leftrightarrow (x \in E \quad y \quad x \in F).$$

Cuando la intersección de dos conjuntos E y F no es vacía, se dice que "E *y* F *se cortan*" o que "E *corta* F" o que "F *corta* E" y también que "E *y* F *se encuentran*".

Cuando la intersección de dos conjuntos E y F es vacía, se dice que "E *y* F *son disjuntos*".

Se llama *reunión* de dos conjuntos E, F y se designa E ∪ F, que se lee "E *unión* F", el conjunto de los elementos pertenecientes al menos a uno de los conjuntos E, F, luego

$$x \in E \cup F \Leftrightarrow (x \in E \quad o \quad x \in F).$$

La intersección y la reunión de dos conjuntos se representan por las partes rayadas de los esquemas 3 y 4.

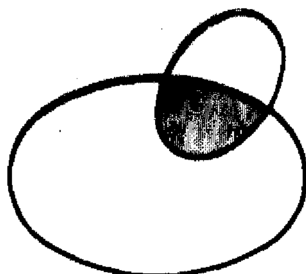


FIG. 3

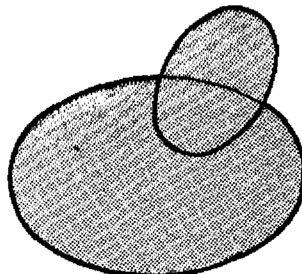


FIG. 4

Se ve inmediatamente que:

la intersección y la reunión son *conmutativas*, es decir, cualesquiera que sean los conjuntos A y B

$$A \cap B = B \cap A \qquad A \cup B = B \cup A;$$

– la intersección y la reunión son *asociativas*, es decir, cualesquiera que sean los conjuntos A, B y C

$$(A \cap B) \cap C = A \cap (B \cap C) \qquad (A \cup B) \cup C = A \cup (B \cup C);$$

– *la intersección es distributiva respecto a la reunión y la reunión distributiva respecto a la intersección*, es decir, para todos los conjuntos A, B y C

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C);$$

en fin, cualquiera que sea A

$$A \cap A = A \qquad A \cup A = A$$
$$A \cap \varnothing = \varnothing \qquad A \cup \varnothing = A.$$

El lector podrá dibujar los esquemas correspondientes a estas distintas igualdades sin que constituyan una demostración, pudiendo, sin embargo, ayudar a construir esta última.

Demostremos, por ejemplo, que

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

sea x un elemento de $A \cap (B \cup C)$, x pertenece a A y a B o C, pertenece, pues, a $A \cap B$ o $A \cap C$;
luego a $(A \cap B) \cup (A \cap C)$.

Sea y un elemento de $(A \cap B) \cup (A \cap C)$, es un elemento de $A \cap B$ o de $A \cap C$, en todos los casos, es un elemento de A y un elemento de B o C, luego y pertenece a A y a $B \cup C$, a $A \cap (B \cup C)$.

La asociatividad de la intersección y de la reunión nos permite definir la intersección de tres, cuatro... conjuntos que designaremos

$$A \cap B \cap C \qquad A \cup B \cup C$$
$$A \cap B \cap C \cap D \qquad A \cup B \cup C \cup D.$$
$$\ldots \qquad \ldots$$

Se puede en particular definir la intersección y la reunión de dos partes A y B de un conjunto E; se demostrará fácilmente que cualesquiera que sean A y B

(1) $$\complement_E (A \cap B) = \left(\complement_E A\right) \cup \left(\complement_E B\right)$$

(1') $$\complement_E (A \cup B) = \left(\complement_E A\right) \cap \left(\complement_E B\right).$$

Se puede generalizar la notación E — A designando por B — A *el conjunto de los elementos pertenecientes a* B *que no pertenecen a* A, luego por definición

$$B - A = B - (A \cap B).$$

En particular si A y B son partes de E

$$B - A = B \cap \complement_E A.$$

EJERCICIOS

1. Se llama *diferencia simétrica* de los conjuntos A y B el conjunto definido por $A \Delta B = (A \cup B) - (A \cap B)$; demostrar que $A \Delta B = (A - B) \cup (B - A)$.

2. Demostrar que si para *toda* parte A de E se tiene
$$A \cup X = E \quad [\text{resp. } A \cap X = A] \quad \text{entonces } X = E.$$

Demostrar que si para *toda* parte A de E se tiene

$$A \cup Y = A \quad [\text{resp. } A \cap Y = \varnothing] \quad \text{entonces } Y = \varnothing.$$

3. Siendo A y B dos partes cualesquiera de E, encontrar todas las partes X tales que $B \cap X = A$, y todas las partes Y tales que $B \cup Y = A$.

4. Demostrar que

$$A \cap B = A \Leftrightarrow A \subset B$$
$$A \cup B = A \Leftrightarrow B \subset A.$$

5. Dadas dos partes propias distintas A y B, expresar todas sus respectivas posiciones.

6. Propiedades definidas sobre un conjunto. Cuantificadores

a) Propiedades definidas sobre E

Dado un conjunto E y una parte A de E, llamaremos *propiedad característica* de los elementos de A *todo criterio* que permita decidir para todo elemento x de E entre las dos proposiciones

$$x \in A, \qquad x \notin A \Leftrightarrow x \in \complement_E A.$$

Pues si p es una propiedad característica de los elementos de A, *no p* es una propiedad característica de los elementos de $\complement_E A$, diremos que p es una *propiedad definida sobre* E.

A "*x pertenece a* A" podemos, pues, sustituir la proposición equivalente "*x posee la propiedad p*", que se escribe abreviadamente "$p(x)$". Así escribiremos

$$A = \{x \in E \mid p(x)\}$$

o bien si no ha lugar a duda

$$A = \{x \mid p(x)\}$$

que se leerá: "A está descrito por los elementos de E poseyendo la propiedad p". Se escribirá, pues,

$$\complement_E A = \{x \in E \mid \text{no } p(x)\}.$$

Por ejemplo, el conjunto de los números reales positivos $\mathbf{R}_+$ estará definido por

$$\mathbf{R}_+ = \{x \in \mathbf{R} \mid x \geq 0\}$$

el conjunto de los números reales estrictamente positivos $\mathbf{R}_+^*$ se designará por

$$\mathbf{R}_+^* = \{x \in \mathbf{R} \mid x > 0\}.$$

Dos propiedades características de los elementos de una misma parte A *de* E *se les llama equivalentes sobre* E.

Por ejemplo, sobre el conjunto E de los *triángulos* T la propiedad p «T tiene dos lados iguales» y q «T tiene dos ángulos iguales» son dos propiedades equivalentes.

b) **Cuantificadores**

Sea p una propiedad definida sobre E, y A una parte de E cuyos elementos tienen como propiedad característica p. Se pueden presentar *tres casos:*

A *no es vacío,* existe, pues, al menos un elemento de E poseyendo p, escribiremos

$$(1) \qquad (\exists\, x \in \mathrm{E})p(x)$$

que se lee: "*existe al menos un elemento x de* E *poseyendo* p"[4].

-- A *es vacío*: ningún elemento de E posee p; se podría escribir: *no* $[(\exists\, x \in \mathrm{E})p(x)]$; más adelante veremos una manera equivalente de escribir esta fórmula.

A *es la parte plena de* E, *todo* elemento de E posee p, se escribirá

$$(2) \qquad (\forall\, x \in \mathrm{E})p(x)$$

que se lee "*cualquiera que sea x de* E, *x posee p*", o más brevemente "*para todo x de* E, $p(x)$".

Los símbolos ∃, ∀ se llaman los *cuantificadores*; no son solamente signos estenográficos para las expresiones "*existe al menos uno*" y "*cualquiera que sea*", sino *símbolos sometidos a reglas de empleo estrictas que derivan de su significado y de los que daremos algunos ejemplos.*

Las fórmulas (1) y (2) traducen los enunciados verdaderos o falsos, independientemente del valor atribuido a x, son las *aserciones* (ver § 2); se podría reemplazar x por cualquier otra letra y, α, ...; se dice que las proposiciones expresadas por *estas fórmulas no contienen a x,* o bien que *x es una variable*[5].

Repitamos que *los cuantificadores definidos en este párrafo se refieren a los elementos de un conjunto determinado* E. No indicaremos este conjunto; para abreviar la escritura, solamente cuando no pueda surgir *ninguna ambigüedad* sobre el conjunto E, se escribirá entonces

$$(\exists\, x)p(x), \qquad (\forall\, x)p(x).$$

(4) Se observará que pueden existir varios elementos de E poseyendo p; si existe *uno y sólo uno* ciertos autores escriben

$$(\exists !\, x \in \mathrm{E})p(x).$$

(5) Esta situación se aproxima a estas que se encuentran en las notaciones siguientes:

$$\bigcup_{i=1}^{n} \mathrm{A}_i, \quad \sum_{i=1}^{n} a_i, \quad \prod_{i=1}^{n} \alpha^{t}a_i, \quad \sum_{i=1}^{n} a_i, \quad \int_a^b f(t)\,dt$$

(ver §§ 32, 33, 34 y 35 para las tres primeras, § 133 *d*, nota 1, para la cuarta y un curso de análisis para la última)

En el ejemplo siguiente esta práctica sería catastrófica: sea p una propiedad característica de una parte propia A de un conjunto E, tenemos a la vez

$$(\exists\, x \in \mathrm{E})p(x), \qquad (\forall\, x \in \mathrm{A})p(x).$$

La introducción de los cuantificadores permite distinguir netamente lo que se llama en la enseñanza elemental «igualdad» e «identidad», por ejemplo,

$$(\forall\, x \in \mathbf{R}) \qquad (x+1)^2 = x^2 + 2x + 1$$
$$(\exists\, x \in \mathbf{R}) \qquad 2x + 1 = 0$$

igualmente

$$[(\forall\, x \in \mathbf{R}) \qquad ax + b = 0] \Leftrightarrow a = b = 0$$
$$[(\exists\, x \in \mathbf{R}) \qquad ax + b = 0] \Leftrightarrow a \neq 0 \quad \text{o} \quad a = b = 0.$$

c) **Relaciones entre los cuantificadores $\exists$ y $\forall$**

Todos los elementos considerados se les supone pertenecientes a E. Las fórmulas

$$(\forall\, x)p(x) \qquad (\exists\, x)p(x)$$

son proposiciones (de hecho las proposiciones particulares que hemos llamado aserciones, ver § 2); busquemos su negación; tenemos

$$(\forall\, x)p(x) \Leftrightarrow \mathrm{A} = \{x \in \mathrm{E} \mid p(x)\} = \mathrm{E}$$

pues $\complement_{\mathrm{E}} \mathrm{A} = \varnothing$. La negación de esta proposición es $\complement_{\mathrm{E}} \mathrm{A} \neq \varnothing$, es decir, "existe al menos un x de E que posee *no* p". Igualmente

$$(\exists\, x)p(x) \Leftrightarrow \mathrm{B} = \{x \in \mathrm{E} \mid p(x)\} \neq \varnothing$$

la negación de esta proposición es $\mathrm{B} = \varnothing$ luego $\complement_{\mathrm{E}} \mathrm{B} = \mathrm{E}$, es decir, "todo elemento de E posee *no* p".

Tenemos, pues,

$$\textit{no}\ [(\forall\, x)p(x)] \Leftrightarrow [(\exists\, x)\ \textit{no}\ p(x)]$$
$$\textit{no}\ [(\exists\, x)p(x)] \Leftrightarrow [(\forall\, x)\ \textit{no}\ p(x)].$$

Estas dos reglas son muy importantes para transformar la negación de una proposición haciendo intervenir una propiedad, en la afirmación de una proposición haciendo intervenir la negación de la propiedad. Precisan, además, el lenguaje corriente que es muchas veces ambiguo; en este último la locución "todos los x no poseen p" podría ser interpretado por "todo x que no posee p" o por "todo x que posee no p".

d) **Relaciones entre partes de un conjunto y propiedades definidas sobre un conjunto**

Sean dos partes A y B de E, p una propiedad característica de los elementos de A y q una propiedad característica de los elementos de B. Como hemos dicho anteriormente "$p(x)$" y "$q(x)$" son propiedades relativas a los elementos de E. La proposición

$$p(x) \Rightarrow q(x)$$

se leerá "todo elemento de E que posee *p*, posee *q*", dicho de otra manera

$$A \subset B$$

de lo que resulta

$$[(\forall x \in E)\,[p(x) \Leftrightarrow q(x)]] \Leftrightarrow [A = B]$$

lo que justifica el término de *propiedades equivalentes* sobre E dadas a *p* y *q* en este caso: las proposiciones correspondientes $p(x)$ y $q(x)$ (relativas a los elementos de E) son proposiciones equivalentes.

Tenemos igualmente

$$A \cap B = \{x \in E \mid p(x) \quad \text{y} \quad q(x)\}$$
$$A \cup B = \{x \in E \mid p(x) \quad \text{o} \quad q(x)\}.$$

Podemos construir el cuadro siguiente dando las correspondencias entre las partes de un conjunto E y las propiedades definidas sobre este conjunto E:

$x \in A$	$p(x)$
$x \in B$	$q(x)$
$A \subset B$	$(\forall x \in E)\,[p(x) \Rightarrow q(x)]$
$A = B$	$(\forall x \in E)\,[p(x) \Leftrightarrow q(x)]$
$x \in \complement_E A$	*no* $p(x)$
$\complement_E\left(\complement_E A\right) = A$	$(\forall x \in E)\,[\textit{no}\ (\textit{no}\ p(x)) \Leftrightarrow p(x)]$
$x \in A \cap B$	$p(x)$ y $q(x)$
$x \in A \cup B$	$p(x)$ o $q(x)$
$A \cap B = \varnothing$	$(\forall x \in E)\,[p(x),\ q(x)$ incompatibles$]$
$A \cap B = \varnothing$ y $A \cup B = E$	$(\forall x \in E)\,[q(x) \Leftrightarrow \textit{no}\ p(x)]$

Se podría prolongar aún este cuadro (transitividad de la inclusión y de la implicación, distributividad de la intersección con relación a la reunión, distributividad de la conjunción en relación con la disjunción, etc.); en particular las reglas de dualidad enunciadas en el § 2 corresponden a las fórmulas (1) y (1') del § 5.

Estas dos maneras de razonar, ya sobre los *elementos* de E, ya sobre las *propiedades* de los elementos de E, son rigurosamente equivalentes. Hasta el último tercio del siglo XIX los *matemáticos preferían razonar en términos de propiedades más que en términos de* elementos y conjuntos, por el reparo en emplear los conjuntos que son con frecuencia infinitos (**N**, conjunto de los puntos de un plano, etc.). Utilizaban, por otra parte, una noción de propiedad extremadamente vaga, que, junto con la ambigüedad del verbo *ser* (ver § 26) y a la de «*o*», entrañaban muchas veces errores en los razonamientos, en particular en la negación de las proposiciones. La definición precisa que hemos dado

de la noción de *propiedad definida sobre un conjunto* a partir de la noción de *parte de un conjunto* y el uso de los cuantificadores hacen desaparecer estas dificultades.

Esta correspondencia biunívoca entre las partes de un conjunto E, y las propiedades definidas sobre E, muestran la equivalencia de dos puntos de vista clásicos de la lógica tradicional: punto de vista de la extensión (razonamientos sobre los individuos), punto de vista de la comprensión (razonamiento sobre las propiedades): los matemáticos modernos han adoptado claramente el punto de vista de la extensión.

7. Generalización de la noción de reunión de intersección

Consideremos un conjunto E y un conjunto $\mathfrak{F}$ de partes E, se llama también una *familia* $\mathfrak{F}$ de partes de E; $\mathfrak{F}$ es, pues, una parte de $\mathfrak{P}(E)$.

Se llamará *intersección de la familia* $\mathfrak{F}$ que se escribirá

$$\bigcap_{F \in \mathfrak{F}} F$$

la parte de E tal que cada uno de sus elementos pertenece a *todos* los conjuntos de $\mathfrak{F}$, luego

$$\bigcap_{F \in \mathfrak{F}} F = \{x \mid (\forall F \in \mathfrak{F})\, x \in F\}.$$

Se llamará *reunión de la familia* $\mathfrak{F}$, que se designará

$$\bigcup_{F \in \mathfrak{F}} F$$

la parte de E tal que cada uno de sus elementos pertenece *al menos a uno* de los conjuntos de $\mathfrak{F}$, luego

$$\bigcup_{F \in \mathfrak{F}} F = \{x \mid (\exists F \in \mathfrak{F})\, x \in F\}.$$

Algunas familias de partes de un conjunto E tienen un papel importante: los recubrimientos y las particiones.

DEFINICIÓN 1. — Un recubrimiento $\mathfrak{R}$ *de una parte* A *de* E *es una familia de partes de* E *cuya reunión contiene* A.

Esta definición puede, pues, expresarse de la manera siguiente

$$(\forall x \in A) \qquad (\exists X \in \mathfrak{R}) \qquad x \in X.$$

DEFINICIÓN 2. — Una partición $\mathfrak{P}$ *de un conjunto* E *es un cubrimiento de* E *cuyos elementos (partes de* E*) son no vacíos y dos a dos disjuntos.*

De donde $\bigcup_{X \in \mathfrak{P}} X = E$ con

$$(\forall X \in \mathcal{P})\ X \neq \varnothing \quad \text{y} \quad (\forall X, X' \in \mathcal{P})(X \neq X' \Longrightarrow X \cap X' = \varnothing)$$

de lo que se deduce: *todo x de* E *pertenece a uno y sólo un elemento de* $\mathfrak{P}$.

III. Producto cartesiano. Relaciones. Correspondencias

8. Producto cartesiano

Dados dos conjuntos A y B descritos, respectivamente, por el elemento x y por el elemento y, se llama *par* (x, y) o *doblete* un objeto tal que

$$(x, y) = (x', y') \Leftrightarrow (x = x' \quad \text{e} \quad y = y'),$$

de donde por negación (ver § 2)

$$(x, y) \neq (x', y') \Leftrightarrow (x \neq x' \quad \text{o} \quad y \neq y')$$

x es la *primera coordenada* e y la *segunda coordenada* del par (x, y).

Los pares (x, y) describen un nuevo conjunto llamado *producto cartesiano* de A y de B designado $A \times B$

$$A \times B = \{(x, y) \mid x \in A,\ y \in B\}$$

$A \times B$ se puede leer "A *cruz* B".

ATENCIÓN: No confundir la noción del *par* (x, y) y la noción de *conjunto de dos elementos* $\{x, y\}$

$$(x, y) \in A \times B \qquad \{x, y\} \in \mathfrak{P}(A \cup B)$$

en particular $\{x, y\} = \{y, x\}$.

Por el contrario, si $x \in A$, $y \in B$, (x, y) es un elemento de $A \times B$ e (y, x) es un elemento de $B \times A$ y no de $A \times B$ en general, igualmente en el caso en que $A = B$, (x, y) e (y, x) son dos elementos de $A \times B = A \times A$, que son en general distintos (sólo son iguales cuando $x = y$). En este último caso $(A = B)$ se puede escribir $A \times A = A^2$, si no ha lugar a confusión (ver § 50).

Observemos las siguientes propiedades

$$(A' \subset A \quad \text{y} \quad B' \subset B) \Rightarrow A' \times B' \subset A \times B$$
$$A \times B = \varnothing \Leftrightarrow (A = \varnothing \quad \text{o} \quad B = \varnothing)$$
$$A \times (B \cap C) = (A \times B) \cap (A \times C)$$
$$A \times (B \cup C) = (A \times B) \cup (A \times C)$$
$$(C \neq \varnothing \quad \text{y} \quad A \times C = B \times C) \Rightarrow A = B.$$

Dados tres conjuntos A, B, C, descritos, respectivamente, por x, y y z, se llama *triplete* (x, y, z) tanto al objeto $((x, y), z)$ como al objeto $(x, (y, z))$. Los tripletes (x, y, z) designan un nuevo conjunto llamado producto cartesiano de A, B, C y denotado $A \times B \times C$. Tenemos por convenio

$$A \times B \times C = (A \times B) \times C = A \times (B \times C).$$

Si $A = B = C$, se puede escribir, si no ha lugar a confusión,

$$A \times A \times A = A^3.$$

EJEMPLOS

1. El plano de la geometría analítica es el conjunto producto $\mathbf{R}^2$; el espacio es el conjunto de los tripletes (x, y, z) de tres números reales, luego el conjunto $\mathbf{R}^3$.

2. El conjunto de las homotecias del espacio, es decir, el conjunto de los pares (A, k), en donde A es un punto del espacio $\mathbf{R}^3$ y k un número real no nulo, luego un elemento de $\mathbf{R}^*$, es el conjunto producto $\mathbf{R}^3 \times \mathbf{R}^*$.

3. El conjunto de las fracciones a/b es el conjunto de los pares (a, b), donde a es un entero cualquiera y b un entero no nulo; el conjunto es, pues, $\mathbf{Z} \times \mathbf{Z}^*$.

6. Relación. Gráfica de una relación

a) Se llama *relación* R entre x e y describiendo, respectivamente, dos conjuntos A y B, *toda propiedad definida sobre* A × B, es decir, *una propiedad característica de los elementos de una parte* G *de* A × B. G se llama *gráfica de la relación* R.

A "*(x, y) pertenece a* G" podemos, pues, sustituirla por la proposición equivalente "*x e y verifican la relación* R" que se designa abreviadamente R(x, y); así

$$R(x, y) \Leftrightarrow (x, y) \in G,$$

o bien

$$G = \{(x, y) \in A \times B \mid R(x, y)\}.$$

Igualmente

$$\text{no } R(x, y) \Leftrightarrow (x, y) \in \complement_{A \times B} G$$

y de un modo más general todo lo que hemos dicho de las propiedades definidas sobre E es válido para las relaciones entre elementos de A y de B, puesto que son las propiedades definidas sobre A × B. En particular si R y R′ tienen por gráficas respectivas G y G′

$$G \subset G' \Leftrightarrow [R(x, y) \Rightarrow R'(x, y)]$$
$$G = G' \Leftrightarrow [R(x, y) \Leftrightarrow R'(x, y)].$$

En este último caso se dice que las relaciones R y R′ son *equivalentes.*

b) Alguna vez se consideran fórmulas como

$$(\forall x \in A) \qquad (\exists y \in B) \qquad R(x, y)$$

que se enuncia: "*Para todo x de* A, *existe un y de* B *tal que x e y verifican la relación* R". *El orden de los cuantificadores es esencial.* Escribamos todas las fórmulas posibles (siendo x un elemento de A e y de B, tomaremos la notación abreviada)

(1)	$(\forall x)$	$(\forall y)$	$R(x, y)$	(1′)	$(\forall y)$	$(\forall x)$	$R(x, y)$
(2)	$(\exists x)$	$(\exists y)$	$R(x, y)$	(2′)	$(\exists y)$	$(\exists x)$	$R(x, y)$
(3)	$(\forall x)$	$(\exists y)$	$R(x, y)$	(4)	$(\exists y)$	$(\forall x)$	$R(x, y)$
(5)	$(\exists x)$	$(\forall y)$	$R(x, y)$	(6)	$(\forall y)$	$(\exists x)$	$R(x, y)$.

(1) y (1') son equivalentes, igualmente (2) y (2')[6]. Contrariamente no es lo mismo para (3) y (4); tomemos, por ejemplo, A = B = **N**, siendo la relación R, $x \leq y$

$$(\forall x) \quad (\exists y) \quad x \leq y \qquad \text{es verdadera}$$
$$(\exists y) \quad (\forall x) \quad x \leq y \qquad \text{es falsa.}$$

Se podrá demostrar a título de ejercicio que (5) entraña (6).

Observemos que una vez los conjuntos A y B y la relación R escogidos, las fórmulas precedentes no contienen ni x, ni y: cada una de ellas es o bien verdadera, o bien falsa, son aserciones. En el mismo orden de ideas, la fórmula

$$(\exists y) \quad R(x, y)$$

no contiene y (ver § 6, *b*), los conjuntos A, B y la relación R escogidos, esta proposición será o bien verdadera, o bien falsa siguiendo la elección de x en A: es, pues, una propiedad de x; igualmente

$$(\forall x) \quad R(x, y)$$

es una propiedad de y, etc. Esta observación nos permitirá formar la negación de una aserción tal como (1), (1'), ..., (6). Consideremos (3), por ejemplo, pongamos

$$p(x) \Leftrightarrow (\exists y) \quad R(x, y)$$

obtendremos las afirmaciones siguientes, todas equivalentes entre sí

$$\text{no } [(\forall x) \quad (\exists y) \quad R(x, y)] \Leftrightarrow \text{no } [(\forall x) \quad p(x)]$$
$$(\exists x) \text{ no } p(x) \Leftrightarrow (\exists x) \text{ no } [(\exists y) \quad R(x, y)]$$
$$(\exists x) \quad (\forall y) \text{ no } R(x, y).$$

Así, pues, utilizando dos veces el resultado de **6, c** *vemos que se obtiene la negación de* (3) *reemplazando cada* ∀ *por un* ∃ *y cada* ∃ *por un* ∀, *y reemplazando* R(x, y) *por no* R(x, y).

Se comprobará fácilmente que la regla es general; en particular es esta regla la que realza el interés del uso de los cuantificadores.

c) Se definirá una *relación* R(x, y, z) entre x elemento de A, y elemento de B, z elemento de C; como aquella que expresa una propiedad definida sobre A × B × C, es decir, es una propiedad característica de los elementos de una parte G de A × B × C, G es el *gráfico* de la relación R.

(6) Las fórmulas (1) y (2) podrían escribirse, respectivamente, si A = B = E

$$(\forall (x, y) \in E^2) \quad R(x, y), \qquad (\exists (x, y) \in E^2) \quad R(x, y)$$

se utilizan a menudo las fórmulas abreviadas

$$(\forall x, y \in E) \quad R(x, y), \qquad (\exists x, y \in E) \quad R(x, y)$$

o si no ha lugar a confusión

$$(\forall x, y) \quad R(x, y), \qquad (\exists x, y) \quad R(x, y).$$

Una fórmula como la

$$(\forall x) \quad (\exists y) \quad (\forall z) \quad R(x, y, z)$$

es una aserción, cuya negación es

$$(\exists x) \quad (\forall y) \quad (\exists z) \text{ no } R(x, y, z).$$

d) Estudiaremos en fin el caso particular en que A = B = E; una **relación** entre x e y de E se llama *relación binaria entre elementos de* E o *relación binaria definida* sobre E; está caracterizada por su *grafo* o *gráfica* G, que es una parte de E × E.

Por ejemplo, la igualdad $x = y$ es una relación binaria definida sobre un conjunto E cualquiera, su grafo Δ se llama la *diagonal* de E × E, Δ está descrito por los pares (x, x), describiendo x el conjunto E.

Si una relación definida sobre E es verdadera para todo par (x, y), se dice algunas veces que es una *identidad*, su grafo es E × E.

Una relación binaria definida sobre E es

— *reflexiva* si $(\forall x) \quad R(x, x)$
— *simétrica* si $(\forall x, y) \; [R(x, y) \implies R(y, x)]$
— *antisimétrica* si $(\forall x, y) \; [R(x, y) \text{ y } R(y, x)] \iff x = y$
— *transitiva* si $(\forall x, y, z) \; [R(x, y) \text{ y } R(y, z)] \implies R(x, z).$

Ejercicios

1. Caracterizar las propiedades de los grafos de las relaciones binarias definidas sobre E que tengan una de las propiedades precedentes (reflexividad, simetría, etc.).

2. Se consideran las relaciones siguientes definidas sobre N

$$x \leq y, \quad x = y - p, \quad x + y = p \quad (p \text{ entero natural dado}).$$

Determinar las propiedades de estas distintas relaciones, así como su grafo.

3. Iguales preguntas para cada una de las relaciones definidas sobre E = {1, 2, ..., 10}: x divide a y, x e y primos entre sí.

10. Correspondencia entre elementos de un conjunto A y elementos de un conjunto B

a) Sea R una relación entre x elemento de A e y elemento de B, sea G su grafo; se llama *correspondencia* entre A y B el *triplete* (A, B, G). A es el *conjunto de salida*, B el *conjunto de llegada*, G el *grafo de la correspondencia*. Se llama *corte según* x el conjunto de los elementos y de B tales que los pares (x, y) pertenecen a G, igualmente *corte según* y el conjunto de los elementos x de A tales que los pares (x, y) pertenecen a G.

El *conjunto de definición* de la correspondencia (A, B, G) es el conjunto X de los elementos x de A tales que los cortes según x no sean vacíos; el *con-*

junto de los valores de la correspondencia es el conjunto Y de los elementos y de B tales que los cortes según y no sean vacíos. Luego

$$X = \{x \mid (\exists y \in B) \quad (x, y) \in G\}$$
$$Y = \{y \mid (\exists x \in A) \quad (x, y) \in G\}.$$

Se dice también que para *todo x* de X *la correspondencia está definida* y que todo Y es un *valor tomado por la correspondencia.*

Estos diferentes conjuntos están "figurados" en la figura 5, el corte según $x = a$ es no vacío, igualmente el corte según $y = b$, los cortes según $x = a'$ e $y = b'$ son vacíos.

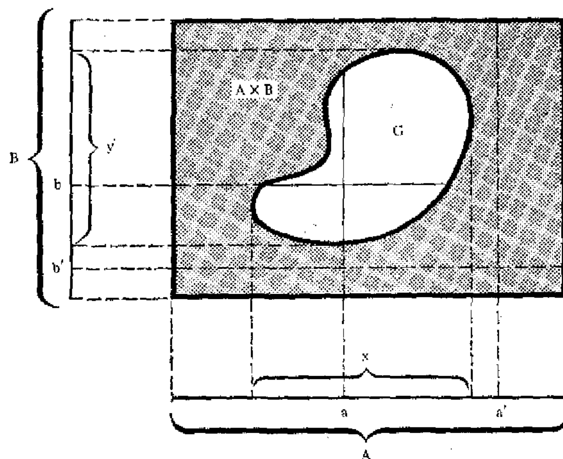


Fig. 5

b) Una correspondencia da origen a dos familias *"de ecuaciones"*

$$(1) \quad R(x, b) \qquad (2) \quad R(a, y).$$

Resolver (1) es buscar el corte según $y = b$, los elementos de este corte (elementos de A) son las *soluciones* de (1). Lo mismo para (2) las soluciones son los elementos del corte según $x = a$.

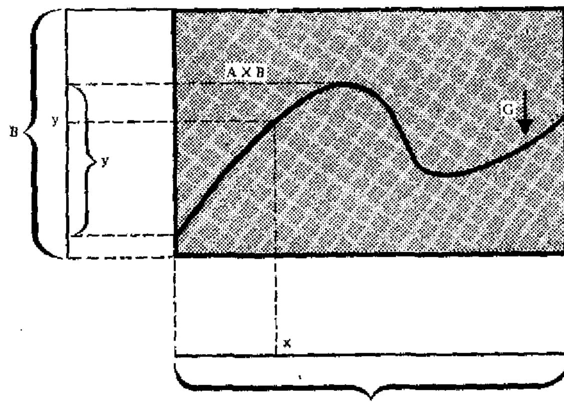


Fig. 6

c) Una correspondencia (**A**, **B**, **G**) es *funcional en relación con la segunda variable* si, cualquiera que sea x de **A**, existe un elemento *único* y ∈ **B** tal que *(x, y)* ∈ **G**. Dicho de otra manera, *el conjunto de definición es idéntico al conjunto de salida y todos los cortes según x contienen un elemento único.* La siguiente sección se va a dedicar al estudio de estas correspondencias particulares y muy importantes.

Cuando la correspondencia es funcional respecto a las dos variables:

Para todo $x \in$ **A**, existe y única de **B**, tal que $(x, y) \in$ **G**.

Para todo $y \in$ **B**, existe x única de **A**, tal que $(x, y) \in$ **G**.

Se dice que es una correspondencia biunívoca entre **A** y **B**.

EJERCICIO

A = B = **R** se considera las correspondencias definidas por las relaciones

$$x^2 = y, \quad \text{sen } x = y, \quad x^2 \leq y, \quad x^3 = y, \quad x^2 + y^2 = 1$$

determinar para cada una de ellas los cortes según $x = a$, según $y = b$, su conjunto de definición, su conjunto de valores. Determinar las que son funcionales en x, las que son funcionales en y.

IV. Aplicación de A en B

1. Noción de aplicación (o de función)

Dados dos conjuntos A y B, una *aplicación f de* A *en* B *es una correspondencia entre un elemento de* A *y un elemento de* B, *funcional con relación a este elemento de* B.

Dicho de otra manera:

Cualquiera que sea el elemento x de A *la aplicación f hace corresponder a x un elemento único y de* B. *Se dice que f aplica* A *en* B *o también que f es una aplicación de* A *en* B.

La palabra función es sinónima[7] *de la palabra aplicación; se dice que la función f está definida en* A *y toma sus valores en* B.

A *es el conjunto de salida o conjunto de definición de f,* **B** *el conjunto de llegada de f.*

x elemento arbitrario de A es la *variable* o el *argumento* de la función. El elemento único y de B que corresponde a x se designa $f(x)$; es el *valor* de la función en x o la *imagen* de x por f, se lee "f de x".

La *gráfica* de la aplicación o de la función f es la parte A × B definida por

$$G = \{(x, y) \in A \times B \mid y = f(x)\}.$$

Notaciones. — Se escribe

$$f\colon A \to B \qquad \text{o} \qquad A \xrightarrow{f} B$$

(7) El uso hace que se emplee principalmente la palabra «función» cuando el conjunto de llegada es un conjunto de «números», es decir, una parte de **R** o **C**, pero esta regla no es general.

que se lee: "*f aplica* A *en* B". Se escribe igualmente

$$(\forall x \in A) \quad x \to y = f(x) \in B$$

o simplemente

$$x \to y \qquad \text{o} \qquad x \to f(x)$$

cuando no da lugar a confusión; se lee "*x da y por f*" o "*x tiene por imagen y por f*" o "*f envíe x sobre y*".

En lugar de $f(x)$ se escribe en ciertos casos f_x, que se lee "*f índice x*"; volveremos a tratar esta *cuestión de índices* en § 17.

Una aplicación es, pues, un *triplete* $f = (A, B, G)$, dos aplicaciones $f = (A, B, G)$ y $f' = (A', B', G')$ son, pues, iguales si y sólo si

$$A = A' \qquad B = B' \qquad G = G',$$

es decir, si tienen el *mismo conjunto de salida, igual conjunto de llegada* y si

$$(\forall x \in A) \qquad f(x) = g(x)$$

se escribe entonces $f = g$.

Serán *desiguales* ($f \neq g$) si al menos una de estas condiciones no se cumple. En particular si $A = A'$ y $B = B'$, $f \neq g$ es equivalente a

$$(\exists x \in A) \qquad f(x) \neq g(x).$$

El conjunto de todas las aplicaciones de A en B es un nuevo conjunto designado $\mathfrak{F}(A, B)$.

OBSERVACION

Cuando no ha lugar a confusión hemos indicado que se puede decir «*sea la aplicación (o la función)* $f : x \mapsto f(x)$»; por el contrario, la expresión «*sea la función* $f(x)$» es un *abuso de lenguaje grave*; en efecto,

$$f(x) \in B \qquad f \in \mathfrak{F}(A, B).$$

Esta confusión entre el valor de una función en x y la función f, demasiado frecuente, es la causa de múltiples errores.

Así no se debe decir «*la función* cos x», sino
— la función $x \mapsto \cos x$;
— o la función *coseno*.

12. Ejemplos

a) $A = B = \mathbf{R}$ (a, b, c reales).

$$x \to ax + b, \qquad x \to ax^2 + bx + c, \qquad x \to \operatorname{sen} x$$
$$x \to f(x) \quad \text{con} \quad f(x) = 0 \quad \text{si} \quad x \in \mathbf{Q} \quad \text{y} \quad f(x) = 1 \quad \text{si} \quad x \notin \mathbf{Q}$$

son funciones definidas en $\mathbf{R}$ y con valores en $\mathbf{R}$.

$A = [-1, +1] \subset \mathbf{R}$, $B = \mathbf{R}$, $x \to \sqrt{1-x^2}$ es una función definida en $[-1, +1]$ y con valores en $\mathbf{R}$.

De una manera general se llama *función numérica real* toda aplicación de un conjunto E en **R**, y *función numérica real de variable real* toda aplicación de una parte de **R** en **R**.

b) Las "*transformaciones geométricas* punto a punto del espacio" son aplicaciones del conjunto $\mathbf{R}^3$ en sí mismo; por ejemplo, las traslaciones, rotaciones, homotecias de razón $k \neq 0$, etc.

La inversión $I(0, k \neq 0)$ es una aplicación de $\mathbf{R}^3 - \{0\}$ en sí mismo.

c) La aplicación f de A en B definida por

$$(\forall x \in A) \qquad f(x) = b$$

es una *función constante* (donde b es un *elemento particular* de **B**).

La aplicación f de A en A definida por

$$(\forall x \in A) \qquad f(x) = x$$

es la *aplicación idéntica* de A, o *identidad* de A; se le designa id_A.

Si $A \subset B$, la aplicación f de A en B tal que

$$(\forall x \in A) \qquad f(x) = x$$

es la *aplicación canónica de una parte* A *de* B *en* B.

d) Si el conjunto de salida es el producto cartesiano $A \times B$ y el conjunto de llegada es C, la aplicación

$$f: A \times B \to C$$

definido también por

$$(x, y) \to z = f(x, y)$$

se la llama *función de dos variables,* o de dos argumentos x, y.

Así la aplicación de $A \times B$ en A definida por $(x, y) \to x$ es la *función primera coordenada designada* pr_1. Se definirá igualmente pr_2; en consecuencia, se tiene

$$pr_1(x, y) = x \qquad pr_2(x, y) = y.$$

A partir de una aplicación f de $A \times B$ en C, se puede definir una familia de aplicaciones de A en C: supongamos y fijo, la aplicación de A en C definida por

$$x \to f(x, y)$$

depende de y; se le llama *la aplicación parcial definida por f relativa al valor y del segundo argumento*; se le designa f_y o $f(., y)$. Se tiene, pues,

$$f_y : A \to C \quad \text{con} \quad f_y(x) = f(x, y).$$

Se definiría de manera semejante f_x o $f(x, .)$ que es una aplicación de B en C relativa al valor x del primer argumento; se tiene, pues,

$$y \to f_x(y) = f(x, y).$$

e) Sea E un conjunto y A una de sus partes; la aplicación φ_A de E en $\{0, 1\}$ definida por

$$\begin{array}{ll} x \in A & \varphi_A(x) = 1 \\ x \in E - A & \varphi_A(x) = 0 \end{array}$$

se llama *función característica de la parte* A *de* E.

Es una aplicación parcial de la aplicación φ de $E \times \mathfrak{P}(E)$ en $\{0, 1\}$ definida por

$$(x, A) \to \varphi(x, A) = \left\langle \begin{array}{lll} 1 & \text{si} & x \in A \\ 0 & \text{si} & x \in E - A. \end{array} \right.$$

EJERCICIO

Demostrar que, siendo A y B dos partes de E no vacío, para todo x de E

$$\begin{aligned} \varphi_{E-A}(x) &= 1 - \varphi_A(x) \\ \varphi_A(x) = \varphi_B(x) &\Leftrightarrow A = B \\ \varphi_{A \cap B}(x) &= \varphi_A(x)\varphi_B(x) \\ \varphi_{A \cup B}(x) &= \varphi_A(x) + \varphi_B(x) - \varphi_A(x)\varphi_B(x). \end{aligned}$$

13. Imágenes e imágenes recíprocas de subconjuntos

Sea f una aplicación de A en B y X una parte de A, se llama ***imagen de* X *por* f** y se escribe $f(X)$ el subconjunto de B definido por

$$f(X) = \{f(x) \mid x \in X\};$$

dicho de otra manera, $f(X)$ se halla descrito por $f(x)$ cuando x describe X.

En particular, $f(A)$ se llama *la imagen de* f, es un abuso de lenguaje por: la imagen del conjunto de salida de f por f, se la designa algunas veces Im f.

Siendo Y un subconjunto de B, se llama ***imagen recíproca de* Y *por* f** y se escribe $f^{-1}(Y)$ el subconjunto de A definido por

$$f^{-1}(Y) = \{x \mid f(x) \in Y\}.$$

Observemos que $f(X)$ es vacío si y sólo si X es vacío. Por el contrario, $f^{-1}(Y)$ puede ser vacío sin que Y lo sea, por ejemplo, si existe Y no vacío tal que

$$Y \subset B - f(A).$$

Si $X = \{x\}$, $f(X)$ tiene un solo elemento, es el conjunto $\{f(x)\}$, es una parte de B, mientras que $f(x)$ es un elemento de B. Contrariamente, $f^{-1}(\{y\})$ puede ser vacío o contener varios elementos; con un abuso de notación, se le designa $f^{-1}(y)$.

Por ejemplo, sea f: $\mathbf{R} \to \mathbf{R}$, $f(x) = \text{sen } x$

$$f^{-1}(2) = \varnothing, \qquad f^{-1}\left(\frac{1}{2}\right) = \left\{ \frac{\pi}{6} + 2k\pi, \quad \frac{5\pi}{6} + 2k\pi \right\} k \in \mathbf{Z}.$$

Se demostrará a título de ejercicio los resultados siguientes, si X_1 y X_2 son dos partes de A e Y_1 e Y_2 son dos partes de B

(1)	$X_1 \subset X_2 \Rightarrow f(X_1) \subset f(X_2)$	(1′)	$Y_1 \subset Y_2 \Rightarrow f^{-1}(Y_1) \subset f^{-1}(Y_2)$
(2)	$f(X_1 \cup X_2) = f(X_1) \cup f(X_2)$	(2′)	$f^{-1}(Y_1 \cup Y_2) = f^{-1}(Y_1) \cup f^{-1}(Y_2)$
(3)	$f(X_1 \cap X_2) \subset f(X_1) \cap f(X_2)$	(3′)	$f^{-1}(Y_1 \cap Y_2) = f^{-1}(Y_1) \cap f^{-1}(Y_2)$
(4)	$f^{-1}(f(X)) \supset X$	(4′)	$f(f^{-1}(Y)) \subset Y.$

Se observará en particular los resultados (3), (4) y (4′). Para estos tres casos, los ejercicios 2 y 3 del § 14 darán una condición necesaria y suficiente para que la inclusión sea reemplazada, bien por la inclusión estricta, bien por la igualdad.

EJERCICIO

Con la ayuda de la función pr_1: $E_1 \times E_2 \to E_1$ dar un ejemplo en el que

$$f(X_1 \cap X_2) \neq f(X_1) \cap f(X_2).$$

Cuando $A = B$ se pueden producir algunas particularidades:

Si $f(X) \subset X$, se dirá que X es una *parte estable por f*.

Si $f(X) = X$, se dirá que X es una *parte invariante por f*. Un elemento x de A tal que $f(x) = x$ se dice *invariante por f*.

Para una parte X invariante, se distinguirá una *parte X descrita por elementos invariantes* de una parte en la que no todos los elementos son invariantes (*parte globalmente invariante*). Por ejemplo, en la inversión plana $I(O, R^2)$ para el círculo (O, **R**) cada punto es invariante, mientras que todo círculo que es ortogonal al círculo (O, R) es globalmente invariante.

14. Suprayecciones. Inyecciones. Biyecciones. Aplicación recíproca de una biyección

a) Dada una aplicación f de A en B e y un elemento de B, surgen dos preguntas:

1.ª $f^{-1}(y)$, ¿es vacío o no?

2.ª Si $f^{-1}(y)$ no es vacío, este subconjunto de A ¿contiene uno o varios elementos?

Conocemos la respuesta de la primera pregunta (si $A \neq \varnothing$)

$$y \in f(A) \Leftrightarrow f^{-1}(y) \neq \varnothing,$$

lo que nos permite definir una clase particular de aplicación:



DEFINICIÓN 1. — *Una aplicación f de* A *en* B *es* suprayectiva, *o también es una* suprayección *si y sólo si*

$$f(A) = B.$$

Se dice también que f es una aplicación de A sobre B.

Si f es suprayectiva para todo y de B, $f^{-1}(y)$ no es vacía, se dice también que f *transforma* A en B.

En cuanto a la segunda pregunta, para y perteneciente a $f(A)$, nos lleva a definir una clase particular de aplicaciones, aquellas tales para las que cuando $f^{-1}(y)$ es no vacío, contiene uno y sólo un elemento:

DEFINICIÓN 2. — *Una aplicación f de* A *en* B *es inyectiva o también es una* inyección *si y sólo si*

$$(\forall x, x' \in A) \quad [f(x) = f(x') \implies x = x']$$

o también si y sólo si

$$(\forall x, x' \in A) \quad [x \neq x' \implies f(x) \neq f(x')].$$

DEFINICIÓN 3. — *En fin, una aplicación de* A *en* B *es* biyectiva, *o bien es una* biyección *de* A *sobre* B, *si es a la vez suprayectiva e inyectiva. Se dice también que es una* aplicación biunívoca *de* A *sobre* B.

b) A toda aplicación f de A en B y a todo elemento b de B corresponde una *ecuación* (ver § 10 *b*), es decir, la propiedad de x

$$f(x) = b$$

verdadera para ciertos elementos x de A y falsa para los otros. Estos valores de x, que son los elementos de $f^{-1}(b)$ son las *soluciones* de la ecuación; encontrarlas es *resolver* la ecuación.

Las definiciones precedentes nos permiten presentar la discusión siguiente:

f cualquiera : $b \in f(A)$: al menos una solución,
$b \notin f(A)$: ninguna solución.

f suprayectiva : cualquiera que sea b: al menos una solución.

f inyectiva : $b \in f(A)$: una única solución,
$b \notin f(A)$: ninguna solución.

f biyectiva : cualquiera que sea b: una solución única.

TEOREMA. — *La aplicación f de* A *en* B *es biyectiva si y sólo si la ecuación $f(x) = y$ tiene una única solución cualquiera que sea y de* B.

c) Supongamos f biyectiva; cualquiera que sea y la ecuación $y = f(x)$ tiene una solución única; designémosla por $x = g(y)$, definamos así una aplicación de B en A

$$g: B \to A \qquad y \to x = g(y)$$

que es visiblemente biyectiva, pues

$$y = f(x) \Leftrightarrow x = g(y).$$

La aplicación g se llama aplicación recíproca de la aplicación biyectiva f, se la designa f^{-1}, lo que no presenta ningún inconveniente, pues la equivalencia anterior muestra que $g(Y) = f^{-1}(Y)$.

Una biyección y su recíproca pueden representarse

$$A \underset{f^{-1}}{\overset{f}{\rightleftarrows}} B$$

por otro lado,

$$(\forall x \in A)(\forall y \in B) \qquad [y = f(x) \Leftrightarrow x = f^{-1}(y)].$$

OBSERVACIONES

1. Algunos autores llaman *función inversa* a la función recíproca de una biyección. Esta costumbre *se debe desechar*, pues si f es, por ejemplo, una aplicación de A en $\mathbf{R}^*$ la inversa de f es la aplicación definida por $x \to 1/f(x)$.

2. No se confundirán las fórmulas $f^{-1}(Y)$ y $f^{-1}(y)$ utilizadas en todos los casos (imagen recíproca de Y, o de y por f) y la fórmula f^{-1} utilizada únicamente cuando f es biyectiva.

3. Se puede considerar la aplicación F de $\mathfrak{P}(A)$ en $\mathfrak{P}(B)$ definida por

$$X \subset A \qquad F(X) = f(X)$$

que se llama *extensión de f a los conjuntos de las partes de* A *y de* B y la aplicación G de $\mathfrak{P}(B)$ en $\mathfrak{P}(A)$ definida por

$$Y \subset B \qquad G(Y) = f^{-1}(Y).$$

Ciertos autores representan F y f con el mismo símbolo, así como G y f^{-1}; desgraciadamente en general, F y G no son biyectivos (ver filas 4 y 4′ del § 13 y ejercicio 2 a continuación) y F y G no son recíprocas una de otra. Nosotros no haremos esta asimilación y nos limitaremos a las fórmulas definidas en el § 12 y recordadas en la observación 2 más arriba.

EJERCICIOS

1. Entre los ejemplos dados en el parágrafo 11, determinar las aplicaciones que son suprayectivas, inyectivas, biyectivas. Determinar las aplicaciones recíprocas de las biyecciones.

2. Demostrar que

[cualquiera que sea $X \subset A$, $X' = f^{-1}(f(X)) = X] \Leftrightarrow f$ es inyectiva
[cualquiera que sea $Y \subset B$, $Y' = f(f^{-1}(Y)) = Y] \Leftrightarrow f$ es suprayectiva.

Dad ejemplos en que

$$X' \neq X \qquad \text{e} \qquad Y' \neq Y.$$

3. Demostrar que

[cualquiera que sean X_1 y X_2 de A, $f(X_1 \cap X_2) = f(X_1) \cap f(X_2)] \Leftrightarrow f$ es inyectiva.

15. Composición de aplicaciones

a) Sean tres conjuntos A, B, C distintos o no y dos aplicaciones f de A en B y g de B en C definidas por

$$x \rightarrow y = f(x) \qquad y \rightarrow z = g(y)$$

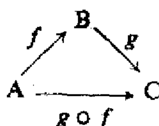
se puede definir una aplicación h de A en C por

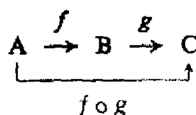
$$(\forall x \in A) \qquad z = h(x) = g(f(x));$$

h es la aplicación compuesta de g y de f y se representa $g \circ f$, luego

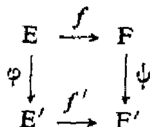
$$(\forall x \in A) \qquad g(f(x)) = (g \circ f)(x).$$

Esta operación, que asocia a toda aplicación f de $\mathfrak{F}(A, B)$ y g de $\mathfrak{F}(B, C)$ una aplicación determinada $g \circ f$ de $\mathfrak{F}(A, C)$, se llama "composición de las aplicaciones"; puede ser representada por uno de los diagramas siguientes:





Dado el diagrama



se dice que este *diagrama es conmutativo* si y sólo si $f' \circ \varphi = \psi \circ f$.

EJEMPLOS

1. Si $A = B = C = \mathbf{R}$, $g \circ f$ es la aplicación llamada «*función de función*» en matemáticas elementales.

2. Si $A = B = C = E$, E espacio de la geometría elemental, f y g dos transformaciones puntuales de E (desplazamientos, homotecias, etc.), $g \circ f$ es el «*producto*» de las dos transformaciones puntuales f y g.

OBSERVACIONES

1. En la notación $g \circ f$, *g escrito primeramente* es la aplicación *efectuada en segundo lugar*, análogamente *f escrita en segundo lugar* es la aplicación *efectuada en primer lugar*, se dirá que g es la *aplicación de izquierda* y no la primera aplicación, que daría lugar a confusión, igualmente f será la *aplicación de derecha*.

2. En lugar de *aplicación compuesta*, se emplea algunas veces (ver ejemplo 2 más arriba) el término de *aplicación producto*; *en general esta expresión no se debe usar*[8].

(8) Se podrá decir en rigor «producto» cuando no haya lugar a confusión; por ejemplo, «producto de dos traslaciones, de dos desplazamientos...» o producto de dos permutaciones de un conjunto (ver § 85).

pues si $A = B = C = \mathbf{R}$, por ejemplo, se llamará aplicación producto (ver § 54) de g y f la aplicación p definida por

$$(\forall x \in \mathbf{R}) \qquad p(x) = g(x)f(x).$$

Así si

$$\begin{array}{llll} (x) & x^2 & g(x) & \text{sen } x \\ h(x) & (g \circ f)(x) & \text{sen } (x^2) & \\ p(x) & x^2 \text{ sen } x & & \end{array}$$

naturalmente $h \neq p$.

b) Sean cuatro conjuntos A, B, C, D y tres aplicaciones definidas por el diagrama siguiente

$$A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D.$$

Comparemos $(h \circ g) \circ f$ y $h \circ (g \circ f)$; observemos antes que

$$h \circ g \in \mathfrak{F}(B, D) \Rightarrow (h \circ g) \circ f \in \mathfrak{F}(A, D)$$
$$g \circ f \in \mathfrak{F}(A, C) \Rightarrow h \circ (g \circ f) \in \mathfrak{F}(A, D).$$

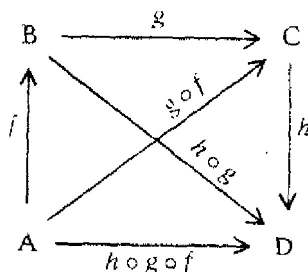
Compararemos los valores de $(h \circ g) \circ f$ y $h \circ (g \circ f)$ para todo valor de A

$$[(h \circ g) \circ f](x) = (h \circ g)[f(x)] = h(g(f(x)))$$
$$[h \circ (g \circ f)](x) = h[(g \circ f)(x)] = h(g(f(x)))$$

luego estas dos aplicaciones, teniendo el mismo conjunto de salida A y el mismo conjunto de llegada D y tomando el mismo valor cualquiera que sea x de A, son iguales

$$(h \circ g) \circ f = h \circ (g \circ f)$$

se expresa esta propiedad diciendo que la composición de las aplicaciones es asociativa; se puede representar esta asociatividad por el diagrama siguiente



y se escribirá simplemente

$$(h \circ g) \circ f = h \circ (g \circ f) = h \circ g \circ f.$$

c) Observemos que para que $g \circ f$ tenga un sentido, *es suficiente* que

$$\text{conjunto de llegada de } f = \text{conjunto de salida de } g,$$

luego en particular si $A = C$, $g \circ f$ y $f \circ g$, aunque tienen sentido al mismo tiempo, éste es tal que

$$g \circ f \in \mathfrak{F}(A, A) \qquad f \circ g \in \mathfrak{F}(B, B),$$

luego finalmente la cuestión relativa a la *conmutatividad* de la composición de las aplicaciones no surgirá, en el caso estudiado, más que cuando $A = B = C$. Un contraejemplo nos mostrará que en general $g \circ f \neq f \circ g$, sea, $A = B = C = \mathbf{R}$ y

$$\begin{array}{ll} f(x) = x^2 & g(x) = \text{sen } x \\ h_1 = g \circ f & h_2 = f \circ g \\ h_1(x) = \text{sen } (x^2) & h_2(x) = (\text{sen } x)^2. \end{array}$$

Sea f una biyección de A en B (B distinto de A) y f^{-1} la aplicación recíproca de f, vemos que

$$A \xrightarrow{f} B \xrightarrow{f^{-1}} A \qquad B \xrightarrow{f^{-1}} A \xrightarrow{f} B,$$

es decir,

$$f^{-1} \circ f = id_A \qquad f \circ f^{-1} = id_B.$$

d) TEOREMA. — *Si f y g son suprayectivas (resp. inyectivas) $g \circ f$ es suprayectiva (resp. inyectiva).*

Si f y g son biyectivas, $g \circ f$ es biyectiva y

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

Escribamos

$$f: A \to B, \qquad g: B \to C, \qquad h = g \circ f: A \to C$$
$$x \to y = f(x) \to z = g(y) = h(x).$$

Supongamos f y g suprayectivas para todo z de C

$$\left\{\begin{array}{ll} (\exists\, y \in B) & z = g(y) \\ & \text{y} \\ (\exists\, x \in A) & y = f(x) \end{array}\right. \Leftrightarrow (\exists\, x \in A) \quad z = h(x)$$

luego h es suprayectiva.

Supongamos f y g inyectivos

$$\begin{array}{ll} h(x) = h(x') \Leftrightarrow g(y) = g(y') \Rightarrow y = y' \\ y = y' \quad \Leftrightarrow f(x) = f(x') \Rightarrow x = x' \end{array}$$

luego h es claramente inyectiva.

Si f y g son biyectivas, análogamente $g\circ f$; cualquiera que sea z de C

$$z = (g\circ f)(x) = g[f(x)] \Rightarrow f(x) = g^{-1}(z) \Rightarrow x = f^{-1}(g^{-1}(z))$$

luego

$$(g\circ f)^{-1} = f^{-1}\circ g^{-1}.$$

e) Siendo f una aplicación de A en A se puede escribir

$$f_1 = f \qquad f_2 = f_1\circ f \dots f_n = f_{n-1}\circ f$$

se dice que f_n es la *iterada n ésima* de f (n entero estrictamente positivo). Si, además, f es biyectiva, se escribirá (n entero estrictamente positivo)

$$f_0 = id_A \qquad f_{-n} = (f^{-1})_n.$$

OBSERVACIÓN

Se escribe algunas veces f^n en lugar de f_n, pero esta notación puede confundirnos en algunos casos; por ejemplo, si A = **R** (ver observación 2 más arriba).

EJERCICIOS

1. Sea $f: A \to B$ y $g: B \to C$; mostrar que $g\circ f$ inyectiva $\Rightarrow f$ inyectiva, y que $g\circ f$ suprayectiva $\Rightarrow g$ suprayectiva.

2. Siendo f, g, h aplicaciones de A en A, mostrar que $f\circ g = f\circ h \Rightarrow g = h$ si y sólo si f es inyectiva, y que $g\circ f = h\circ f \Rightarrow g = h$ si y sólo si f es suprayectiva.

3. Sea $f: A \to B$ y g y h dos aplicaciones de B en A tales que $g\circ f = id_A$, $f\circ h = id_B$, demostrar que f es biyectiva y que $g = h = f^{-1}$.

4. Dados tres conjuntos A, B, C, si existe una biyección f de A sobre B (resp. g de B sobre C) para toda aplicación h de A en C, existe g de B en C (resp. f de A en B) tal que $g\circ f = h$.

5. Teniendo f por espacio de salida A y teniendo g por espacio de salida B, determinar las condiciones necesarias y suficientes verificadas por A, B, $f(A)$, $g(B)$ para que las aplicaciones $g\circ f$ y $f\circ g$ existan simultáneamente.

6. Siendo f una aplicación de A en A demostrar que las propiedades siguientes son equivalentes:

a) f es biyectiva y $f = f^{-1}$;

b) $f\circ f = id_A$.

Se dice entonces que f es *involutiva*.

Entre las aplicaciones siguientes buscar las que son involutivas

$$A = \mathbf{R} \qquad f(x) = ax + b$$

$$A = \mathbf{R} - \left\{ \frac{a}{c}, -\frac{d}{c} \right\} \qquad g(x) = \frac{ax+b}{cx+d} \qquad (c \neq 0).$$

16. Restricción y prolongación de una aplicación

Siendo f una aplicación de A en B y X una parte de A, se llama *restricción* de f a X la aplicación g de X en B definida por

$$(\forall x \in X) \qquad g(x) = f(x)$$

f es una *prolongación* de g.

Se representa algunas veces g por los símbolos f_X o $f \mid X$.

Si la restricción de una función está bien determinada por el conocimiento de f y X, no ocurre lo mismo con las prolongaciones: siendo g conocido, una prolongación f de g tiene sus valores bien determinados en X; por el contrario, en A — X sus valores son arbitrarios.

Siguiendo el mismo orden de ideas, siendo f una aplicación de A en B e Y una parte de B tal que $f(A) \subset Y$, se puede considerar la aplicación h de A en Y definida por $h(x) = f(x)$ para todo x de A, que se le denomina *aplicación con valores en* Y *inducida por* f.

Es frecuente que, por abuso de notación, se representen todas estas aplicaciones por la misma letra f, no sin inconvenientes; consideremos, por ejemplo,

$$f: \mathbf{R} \to \mathbf{R} \qquad f(x) = \text{sen } x$$
$$g: \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to \mathbf{R} \qquad g(x) = \text{sen } x$$
$$h: \mathbf{R} \to [-1, +1] \qquad h(x) = \text{sen } x$$
$$k: \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1, +1] \qquad k(x) = \text{sen } x$$

Estas cuatro funciones no son iguales (sus conjuntos de salida y llegada respectivos no son los mismos); por otra parte, no tienen las mismas propiedades:

f no es ni suprayectiva ni inyectiva.
g es inyectiva y no suprayectiva.
h es suprayectiva y no inyectiva.
k es biyectiva.

Por el contrario, toman las cuatro el mismo valor para todo x de $\left[-\frac{\pi}{2}, +\frac{\pi}{2}\right]$; de una manera general diremos que dos aplicaciones

$$f: A \to B \qquad g: C \to D$$

coinciden sobre una parte común X *de* A *y de* C si

$$(\forall x \in X) \qquad f(x) = g(x),$$

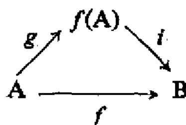
naturalmente esto supone también que $B \cap D$ no sea vacía y contenga $f(X)$.

Las cuatro funciones f, g, h, k coinciden sobre $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, no las confundiremos, pero para simplificar designaremos su valor común en x por sen x.

De manera general, se puede deducir de toda aplicación f de A en B una *aplicación suprayectiva* g *que coincide con* f *sobre* A, la g: $A \to f(A)$ definida por

$$(\forall x \in A) \qquad g(x) = f(x)$$

si, además, f es inyectiva, g será biyectiva. Se observa que si i es la **inyección** canónica de $f(A)$ en B (ver § 12 *b*), se tiene $f = i \circ g$, **fórmula que se puede** representar por el diagrama



f una aplicación cualquiera; g, suprayectiva; i, inyectiva canónica.

En fin, si $A = B$ y si X es una *parte estable* de A por f, es decir, si $f(X) \subset X$, llamaremos *aplicación inducida* sobre X por f, la aplicación g de X en X que coincide con f sobre X.

17. Notación de índices. Familia de elementos. Familia de partes

Hemos dicho que se utilizaba algunas veces en lugar de la notación $f(x)$, la notación por índice f_x (ver § 11). Se dice entonces que la aplicación $f\colon I \to E$ define una *familia de elementos de* E *con índices* pertenecientes al conjunto I que se llama *conjunto de los índices*; en lugar de la notación $i \to a_i$, se emplea a menudo la notación

$$(a_i)_{i \in I}$$

si $J \subset I$ el conjunto $(a_i)_{i \in J}$ es una *subfamilia* de $(a_i)_{i \in I}$; se dice también *familia extraída* de la familia (a_i) $(i \in I)$.

Si el conjunto de los índices es un producto cartesiano $I \times J$, se define una *familia doble*

$$(a_{ij})_{(i,j) \in I \times J} \,.$$

Dado un conjunto E podremos definir la noción de *familia de las partes de* E *con índices*; escribiremos

$$(A_i)_{i \in I}.$$

Hemos definido § 7 la intersección y la reunión de una familia de partes de E, para una familia con índices, tendremos

$$\bigcap_{i \in I} A_i = \{x \mid (\forall\, i \in I) \quad x \in A_i\}$$

$$\bigcup_{i \in I} A_i = \{x \mid \exists\, i \in I) \quad x \in A_i\}.$$

El conjunto de las A_i, cuando i describe I será un *recubrimiento* de F parte de E (ver § 7), si y sólo si

$$F \subset \bigcup_{i \in I} A_i.$$

La aplicación $i \to A_i$ siendo inyectiva, el conjunto de las A_i, recubrimiento de E, cuando i describe I será una *partición* de E si y sólo si (ver § 7)

$$(\forall i \in I) \qquad A_i \neq \varnothing$$

y

$$(\forall i, i' \in I) \quad [i \neq i' \implies A_i \cap A_{i'} = \varnothing],$$

es decir, todo x de E pertenece a una y sólo una parte A_i.

OBSERVACION

Si $i \to f(i) = a_i$ se comete alguna vez el abuso de lenguaje que consiste en llamar familia $(a_i)_{i \in I}$ no a la aplicación f: $I \to E$, sino a la imagen de f, es decir, al conjunto $f(I)$.

V. Relaciones de equivalencia

18. Definición. Clases de equivalencia. Conjunto cociente

a) DEFINICIÓN. — *Una relación binaria* R *entre elementos de un conjunto* E *es una relación de equivalencia si es reflexiva, simétrica, transitiva.*

Se escribirá

$$x \equiv y \pmod{R}$$

que se enuncia "*x e y son equivalentes —o congruentes— módulo* R", o bien "*x es equivalente —o congruente— a y, módulo* R". Luego si R es una relación de equivalencia, tenemos (ver § 9, *d*)

$$(\forall x \in E) \qquad x \equiv x \pmod{R}$$

$$(\forall x, y \in E) \quad [x \equiv y \pmod{R} \implies y \equiv x \pmod{R}]$$

$$(\forall x, y, z \in E) \quad [(x \equiv y \pmod{R} \ \text{ e } \ y \equiv z \pmod{R}) \implies x \equiv z \pmod{R}]$$

b) EJEMPLOS Y EJERCICIOS

Demostrar que las relaciones siguientes son relaciones de equivalencia:

1. En E, $x = y$ (igualdad, o identidad).
2. En E, $x \equiv y$ cualquiera que sean x e y de E (equivalencia absoluta).
3. En **Z**, siendo p un entero estrictamente positivo la relación «$p \mid x - y$», es decir, «p *divide* $x - y$». Se le llama *congruencia módulo p*.
4. En **R** la relación «existe un entero racional k tal que $x - y = 2k\pi$». Se le llama *congruencia módulo* 2π.
5. En el espacio E de la geometría elemental el paralelismo de las rectas (resp. de los planos), considerando por convenio que una recta o un plano es paralela a ella (o a sí mismo).
6. Siendo L el conjunto de los vectores ligados del espacio de la geometría elemental, la *equipolencia*, es decir, la relación «$\overrightarrow{A'B'}$ se deduce de $\overrightarrow{AB}$ por una traslación».
7. Siendo L el conjunto de los vectores ligados del espacio de la geometría elemental, la relación «$\overrightarrow{A'B'}$ se deduce de $\overrightarrow{AB}$ por una traslación paralela a $\overrightarrow{AB}$».
8. Si f es una aplicación de A en B, la relación definida en A por $f(x) = f(x')$.

c) Clases de equivalencia módulo R. Conjunto cociente

Si R es una relación de equivalencia definida sobre E, se llama *clase de equivalencia, módulo* R, toda parte de E descrita por todos los equivalentes a uno de ellos x, se le designará $\dot{x}$, se dirá que x es un *representante* de la clase $\dot{x}$.

Todo x' equivalente a x módulo R pertenece a $\dot{x}$; luego $\dot{x}' \subset \dot{x}$; como la relación $x \equiv x'$ (mod R) es simétrica, se tiene también $\dot{x} \subset \dot{x}'$; finalmente

$$[x \equiv x' \pmod{R}] \Leftrightarrow \dot{x} = \dot{x}'.$$

Consideremos dos clases de equivalencia módulo R, $\dot{x}$ e $\dot{y}$; o bien son disjuntas, o bien existe z de E tal que

$$z \in \dot{x} \qquad y \qquad z \in \dot{y},$$

luego z es equivalente, módulo R, a x y a y y en virtud de la transitividad x e y son equivalentes, módulo R; luego las clases $\dot{x}$ e $\dot{y}$ se confunden; por tanto:

TEOREMA. — *Dos clases de equivalencia, módulo* R, *son disjuntas o confundidas.*

Por otro lado, todo x de E pertenece al menos a una clase $\dot{x}$ (aquella que representa) y a una sola, puesto que dos clases son disjuntas o confundidas; además, ninguna clase es vacía, puesto que $\dot{x}$ contiene al menos a x, luego el conjunto de las clases módulo R es una *partición* de E.

Recíprocamente, sea P una partición de E (ver § 7 y § 17), la relación "x e y *pertenecen al mismo elemento* F *de la partición*" es visiblemente una relación de equivalencia definida sobre E, luego:

TEOREMA Y DEFINICIÓN. — *Dada una relación de equivalencia* R *definida sobre un conjunto* E, *el conjunto de las clases de equivalencia, módulo* R, *es una partición de* E. *Recíprocamente, toda partición de* E *define una relación de equivalencia sobre* E.

El conjunto de las clases de equivalencia de E, *módulo* R, *se llama* conjunto cociente *de* E *por la relación de equivalencia* R *y se designa* E/R.

Siendo las clases de equivalencias partes de E son los elementos de $\mathfrak{P}(E)$, E/R es, pues, una parte de $\mathfrak{P}(E)$, luego un elemento de $\mathfrak{P}(\mathfrak{P}(E))$

$$\begin{array}{lll} x \in \dot{x} & \dot{x} \subset E & \dot{x} \in \mathfrak{P}(E) \\ \dot{x} \in E/R & E/R \subset \mathfrak{P}(E) & E/R \in \mathfrak{P}(\mathfrak{P}(E)). \end{array}$$

Consideremos la aplicación de E en E/R definida por $x \to \dot{x}$; por definición del conjunto E/R es suprayectiva, se le llama la *suprayección canónica* de E sobre E/R.

OBSERVACION

Sobre el término «*representante*» de una clase de equivalencia.

Si $\dot{x}$ comprende más de un elemento, la noción de representante no es una noción canónica. Pero puede ocurrir que haya un *representante privilegiado* determinado por una condición natural:

1. Por ejemplo, en **Z** *para la relación* $p \mid x - y$ $(p > 0)$, para todo x existe q y r únicos (ver § 36) tales que $x = pq + r$ y $0 \leq r < p$ se tiene $\dot{x} = \dot{r}$ $r \in \{0, 1, 2, \dots, p-1\}$ *y* r *es único.*

2. Análogamente en **R** para la relación sen $x =$ sen y existe un x_0 único tal que

$$\dot{x} = \dot{x}_0 \qquad \text{y} \qquad x_0 \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

EJEMPLOS Y EJERCICIOS

Verificar la tabla siguiente dando las clases de equivalencia y los conjuntos cocientes relativos a *los ejemplos anteriores y las denominaciones de algunas de estas clases* y de estos conjuntos cocientes.

Conjunto E	Relación R	Clases módulo R	Conjunto cociente E/R
1. E	$x = y$ (identidad)	$\{x\} \dots$	$\{\{x\}, \dots\} \in \mathfrak{P}(\mathfrak{P}(\mathrm{E}))$
2. E	Equivalencia absoluta	$\mathrm{E} \in \mathfrak{P}(\mathrm{E})$	$\{\mathrm{E}\} \in \mathfrak{P}(\mathfrak{P}(\mathrm{E}))$
3. **Z**	$p \mid x - y$	Enteros módulo p	$\mathbf{Z}/p\mathbf{Z}$
4. **R**	$(\exists\, k \in \mathbf{Z})\ x - y = 2k\pi$	Números reales módulo 2π	$\mathbf{R}/2\pi\mathbf{Z}$
5. Rectas del espacio	Paralelismo	Dirección de recta	Conjunto de todas las direcciones de recta
Planos del espacio	Paralelismo	Dirección de plano	Conjunto de todas las direcciones de plano
6. Vectores ligados del espacio	Equipolencia	Vector libre	Conjunto de los vectores libres
7. Vectores ligados del espacio	$\overrightarrow{\mathrm{A'B'}}$ se deduce de $\overrightarrow{\mathrm{AB}}$ por traslación paralela a $\overrightarrow{\mathrm{AB}}$	*Vector deslizante*	*Conjunto de los vectores deslizantes*
8. E	$f\colon \mathrm{E} \to \mathrm{F}$ $f(x) = f(x')$	$f^{-1}(y)\ (y \in f(\mathrm{E}))$	$\{f^{-1}(y), \dots\}\ y \in f(\mathrm{E})$

Cuando hayamos definido **Z** (estudiaremos con detalle la relación $p \mid x - y$ (ver § 64, *c*)) explicaremos en el § 75 la notación $\mathbf{Z}/p\mathbf{Z}$. Por otro lado, la medida de un ángulo es un número real módulo 2π.

19. Compatibilidad de una relación con una relación de equivalencia

a) Una relación de equivalencia sobre E presenta un mayor interés cuando conserva para x' equivalente a x, módulo R, ciertas de las propiedades de x de donde las siguientes definiciones:

1. Una propiedad P definida sobre E es *compatible* con una relación de equivalencia R definida sobre E si

$$(\forall x, x') \quad [(P(x) \;\; \text{y} \;\; R(x, x')) \Longrightarrow P(x')].$$

2. Una relación binaria S definida sobre E es compatible respecto a x (resp. a y) con una relación de equivalencia R definida sobre E si

$$(\forall x, x', y) \quad [(S(x, y) \;\; \text{y} \;\; R(x, x')) \Longrightarrow S(x', y)]$$

respectivamente,

$$(\forall x, y, y') \quad [(S(x, y) \;\; \text{y} \;\; R(y, y')) \Longrightarrow S(x, y')].$$

3. Una relación binaria S definida sobre E es compatible con una relación de equivalencia R definida sobre E si

$$(\forall x, y, x', y') \quad [(S(x, y) \;\; \text{y} \;\; R(x, x') \;\; \text{y} \;\; R(y, y')) \Longrightarrow S(x', y')].$$

b) En particular, dada una aplicación f de E en F, la relación

$$f(x) = f(x')$$

es una relación de equivalencia definida sobre E y

$$[y = f(x) \qquad \text{y} \qquad f(x) = f(x')] \Rightarrow y = f(x')$$

de donde:

TEOREMA Y DEFINICIÓN. — *Dada la aplicación f de E en F, la relación $y = f(x)$ es compatible respecto a x con la relación de equivalencia $f(x) = f(x')$; esta relación se llama relación de equivalencia asociada a la aplicación f.*

Esto nos permitirá factorizar de una manera canónica toda aplicación. Hemos dicho anteriormente (§ 16) que $f = i \circ g$, siendo g la suprayección de E sobre f(E) que coincide con f sobre E e i la inyección canónica de f(E) en F

$$f(x) = f(x') \Leftrightarrow g(x) = g(x')$$

sea R la relación de equivalencia asociada a f (de E en F) o a g (de E sobre f(E)). Designemos por $\dot{x}$ la clase de equivalencia, módulo R; si se pone $y = f(x) = g(x)$, $\dot{x}$ no es otro que $f^{-1}(y)$, si se considera la suprayección canónica s (ver § 18) definida por $x \to s(x) = \dot{x}$ de E sobre E/R, cada clase $\dot{x} = f^{-1}(y)$ corresponde biunívocamente a y elemento de f(E); dicho de otro modo,

$$g = b \circ s$$

siendo b la biyección de E/R sobre f(E) definido por $\dot{x} \to b(\dot{x}) = y$, tenemos

$$\begin{array}{ccccccc} & s & & b & & i & \\ E & \to & E/R & \to & f(E) & \to & F \\ x & \to & \dot{x} & \to & y & \to & y \end{array}$$

es decir,

$$f = i \circ b \circ s,$$

fórmula que se puede representar por el diagrama

$$\begin{array}{ccc} E & \xrightarrow{f} & F \\ s\downarrow & & \uparrow i \\ E/R & \xrightarrow[b]{} & f(E) \end{array}$$

s suprayección canónica de E sobre E/R.
b biyección canónica de E/R sobre f(E).
i inyección canónica de f(E) en F.

Conclusión. — Las consideraciones desarrolladas en este párrafo y el precedente muestran que, para un conjunto E, definir:

— una partición de E,
— una relación de equivalencia R sobre E,
— una aplicación de E en un conjunto no vacío cualquiera F

son tres aspectos de una sola noción: se trata de descomponer siempre E en subconjuntos no vacíos tales que cada elemento de E pertenezca a uno y sólo a uno de estos subconjuntos.

EJERCICIOS

1. Siendo f una aplicación de E en F, demostrar que hay sobre E relaciones de equivalencia S *distintas de* la $f(x) = f(x')$ *tales que* $y = f(x)$ sea compatible respecto x con S (ver ejercicio número 22, fin del capítulo 1).

2. Determinar las descomposiciones canónicas de las siguientes aplicaciones de **R** en **R**

$$x \to x^2, \qquad x \to \text{sen } x.$$

20. Potencia de los conjuntos

Dados dos conjuntos E y F, consideremos el enunciado *"existe una biyección de* E *sobre* F". Como no existe ningún conjunto del que todo conjunto sea elemento (ver § 3), diremos, por extensión de la noción de relación, que el enunciado precedente define una *relación sobre la clase de todos los conjuntos.*

Precisemos las propiedades de esta relación:

1. Existe una biyección de E sobre E: identidad de E.

2. Si existe una biyección f de E sobre F, existe una biyección de F sobre E; a saber, f^{-1} (ver § 14).

3. Si existe una biyección f de E sobre F y una biyección g de F sobre G, existe una biyección de E sobre G; a saber, $g \circ f$ (ver § 15).

Luego esta relación entre conjuntos cualesquiera posee las tres propiedades de una relación de equivalencia. La relación "*existe una biyección de* E *sobre* F" se traduce por "*los conjuntos* E *y* F *tienen la misma potencia*" o bien "E *y* F *son equipotentes*", y se escribe

$$\text{E eq F.}$$

EJERCICIOS Y EJEMPLOS

1. Todos los conjuntos que tienen n elementos distintos dos a dos son equipotentes entre sí y en particular a $\{1, 2, ..., n-1, n\}$.

2. Los conjuntos equipotentes a **N** o a una parte de **N** se llaman *numerables*. Demostrar que el conjunto de los enteros pares positivos con el 0 es equipotente a **N**, análogamente el conjunto de los enteros positivos con el 0 que son los cuadrados de enteros, o las potencias p-ésimas de enteros (ver § 42).

3. Demostrar que el conjunto de los reales x tales que $0 < x < 1$ es equipotente al conjunto de los reales y tales que $a < y < b$ y que los dos son equipotentes a **R**, se dice que tienen la potencia del *continuo* (utilizar una función lineal, después la función tangente).

VI. Relaciones de orden

21. Relaciones de orden. Conjuntos ordenados. Ejemplos

a) DEFINICIÓN. — *Una relación binaria* R *entre elementos de un conjunto* E *es una relación de orden si es reflexiva, antisimétrica y transitiva.*

En lugar de R(x, y), se designará una relación de orden bien por un signo específico (ver los ejemplos), bien por

$$x \prec y$$

que se enuncia indiferentemente:

- "x es inferior a y", o "y es superior a x".
- "x es anterior a y", o "y es posterior a x".

La relación $y \prec x$ (entre x e y) se llama la *relación de orden opuesto* a $x \prec y$.

Para toda relación de orden tenemos, pues (ver § 9, *d*),

$$(\forall x \in \text{E}) \qquad x \prec x$$
$$(\forall x, y \in \text{E})\ [x \prec y \quad \text{e} \quad y \prec x) \Rightarrow x = y]$$
$$(\forall x, y, z \in \text{E})\ [x \prec y \quad \text{e} \quad y \prec z) \Rightarrow x \prec z].$$

Un conjunto provisto de una relación de orden se llama *conjunto ordenado* por esta relación de orden, se dice también que posee una *estructura de orden*.

Sea A una parte de un conjunto E ordenado por la relación $x \prec y$, para los elementos de A esta relación define una relación de orden sobre A, se dice que esta relación sobre A es *inducida* por la relación $x \prec y$ sobre E.

Representaremos una relación de orden sobre E y las relaciones de orden inducidas sobre las partes de E por el mismo símbolo; una relación de orden definido sobre E es una *prolongación* de la relación de orden definido sobre A.

b) EJEMPLOS Y EJERCICIOS

1. Sobre **R** la relación $a \leq b$ es una relación de orden, ella induce sobre **Q**, **Z**, **N** la relación $a \leq b$. Se enuncia «*a* inferior a *b*» o «*b* superior a *a*».

2. Sobre $\mathbf{N}^*$ la relación $a | b$ que se enuncia «*a divide b*» es una relación de orden.

3. Siendo E un conjunto, sobre $\mathfrak{P}(E)$, la relación $A \subset B$ (inclusión de las partes) es una relación de orden.

4. Sobre el conjunto de los puntos regados por un río, sus afluentes y los afluentes de estos últimos, la relación «A es posterior a B en el sentido de la corriente de las aguas» es una relación de orden.

c) La relación

$$x \prec y \quad \text{y} \quad x \neq y$$

no es una relación de orden: no es ni reflexiva, ni antisimétrica, es solamente transitiva; se enuncia

"*x* es estrictamente inferior a *y*"

o

"*x* es estrictamente anterior a *y*".

OBSERVACION

Algunos autores llaman esta relación «$x \prec y$ y $x \neq y$» relación de orden estricto, *esta expresión no es aconsejable*, pues esta «digamos relación de orden estricto» debería ser un caso particular de relación de orden, luego debería tener al menos las propiedades de una relación de orden; sin embargo, no es así, como acabamos de ver.

En **R** se escribirá $x < y$ (*x estrictamente inferior a y*) por $x \leq y$ y $x \neq y$.
Se verá que nos alejamos del lenguaje al cual, quizás, el lector está acostumbrado enunciando

"*x inferior a y*"	para	"$x \leq y$"
"*x estrictamente inferior a y*"	para	"$x < y$";

para evitar una posible ambigüedad, se puede también decir

"*x inferior a y en sentido amplio*" para "$x \leq y$"

(pero esto recarga el lenguaje).

Así, para $x \geq 0$ (resp. $x \leq 0$), se dirá "*x positivo*" (resp. *x negativo*), "*x positivo en el sentido amplio*" (resp. *x negativo en el sentido amplio*).

22. Orden total, orden parcial. Conjunto totalmente ordenado

a) En el ejemplo 1 (§ 21), cualesquiera que sean *a* y *b* de **R** una de las dos relaciones

$$a \leq b \qquad b \leq a$$

es siempre verdadera. No ocurre lo mismo para los ejemplos 2, 3, 4. En N^* si consideramos 3 y 7, ninguno de los dos divide al otro. De donde las definiciones siguientes:

DEFINICIÓN. — *Se dice que una relación de orden* R *define sobre* E *un* **orden total** *si, cualesquiera que sean a y b, se verifica*

$$a \prec b \quad \text{o} \quad b \prec a.$$

Se dice también que dos elementos cualesquiera son comparables *en el orden definido por* R. *Se dice igualmente que* E *está* totalmente ordenado *por* R *o que* E *posee una* estructura de orden total.

Cuando existe al menos un par (x, y) *de elementos de* E *no comparables por el orden definido por* R, *se dice que* R *define un* orden parcial *o que* E *está* parcialmente ordenado *por* R.

b) En un conjunto totalmente ordenado se llamará *segmento* o *intervalo cerrado* $[a, b]$ e *intervalo abierto* $]a, b[$ (*) las dos partes de E definidas, respectivamente, por

$$[a, b] = \{x \mid a \prec x \prec b\}$$
$$]a, b[= \{x \mid a \prec x \prec b \quad \text{y} \quad x \neq a \quad \text{y} \quad x \neq b\}.$$

Igualmente adaptaremos las notaciones siguientes

(1) $[a, b[= \{x \mid a \prec x \prec b \quad \text{y} \quad x \neq b\}$
(2) $]a, b] = \{x \mid a \prec x \prec b \quad \text{y} \quad x \neq a\}$
(3) $[a, \rightarrow[= \{x \mid a \prec x\}$
(3′) $]a, \rightarrow[= \{x \mid a \prec x \quad \text{y} \quad x \neq a\}$
(4) $]\leftarrow, a] = \{x \mid x \prec a\}$
(4′) $]\leftarrow, a[= \{x \mid x \prec a \quad \text{y} \quad x \neq a\}.$

(1) (resp. (2)) se llama *intervalo semicerrado a la izquierda* (resp. *a la derecha*), *semiabierto a la derecha* (resp. *a la izquierda*).

(3) (resp. (3′)) se llama *sección terminante cerrada* (resp. *abierta*).

(4) (resp. (4′)) se llama *sección principiante cerrada* (resp. *abierta*).

Todas estas notaciones se utilizarán principalmente en **R**, así como en **Q**, **Z**, **N** totalmente ordenados por $x \leq y$.

c) En E totalmente ordenado, se llama *sucesor* a' de a todo elemento tal que $]a, a'[$ sea vacío, se demostrará fácilmente que si tal elemento existe es único; se llamará igualmente *predecesor* $'a$ de a todo elemento tal que $]'a, a[$ sea vacío. Si existe es único.

Por ejemplo, en **R** o **Q** un elemento no tiene ni sucesor ni predecesor; en **Z**: $a' = a + 1$ y $'a = a - 1$.

(*) Con $a \prec b$.

23. Elementos notables de un conjunto ordenado o de una parte de un conjunto ordenado

a) Mayorantes (cotas superiores) y minorantes (cotas inferiores) de una parte X de un conjunto ordenado E

DEFINICIÓN. — *Dada una parte* X *de un conjunto ordenado* E, *un elemento* a *de* E *es un* mayorante (cota superior) *de* X *si*

$$(\forall x \in X) \qquad x \prec a.$$

Se dice que X es una *parte mayorada*(*) de E. Igualmente b de E es un *minorante* de X si

$$(\forall x \in X) \qquad b \prec x.$$

Se dice que X es una *parte minorada*(**) de E.

Una parte a la vez mayorada y minorada se llama una *parte acotada* de E.

Es evidente que todo elemento superior a un mayorante es también un mayorante, y todo elemento inferior a un minorante es también un minorante.

EJEMPLOS

1. En un conjunto totalmente ordenado (por ejemplo, **R**) una sección principiante es mayorada, una sección terminante es minorada, un intervalo es acotado.

2. Designemos por $\mathcal{A}$ el conjunto de las partes de E contenidas en una parte A no vacía de E, $\mathcal{A}$ es mayorada por todas las partes de E que contienen A.

b) Máximo y mínimo elemento de un conjunto ordenado

Supongamos que existe, en E, un M tal que

$$(\forall x \in E) \qquad x \prec M.$$

Este elemento es único; en efecto, sea M′ otro elemento tal que

$$(\forall x \in E) \qquad x \prec M',$$

tendremos a la vez

$$M' \prec M \qquad y \qquad M \prec M',$$

luego $M = M'$; de donde:

TEOREMA Y DEFINICIÓN. — *Si existe en* E *ordenado un elemento superior a todos los demás es único, y se le llama* el máximo elemento *de* E.

Igualmente si existe en E *ordenado un elemento inferior a todos los otros es único, y se le llama* el mínimo elemento *de* E.

Se podrá definir, si existen, el máximo y el mínimo elemento de una parte X de E ordenado; son únicos y pertenecen a X.

(*) Acotada superiormente. (*N. del T.*)
(**) Acotada inferiormente. (*N. del T.*)

EJEMPLOS

1. **R**, **Q**, **Z**, ordenados por $x \leq y$ no tienen ni máximo ni mínimo elemento. N ordenado por $x \leq y$ tiene un mínimo elemento 0.

Un intervalo cerrado $[a, b]$ de **R** tiene un mínimo elemento a, un máximo elemento b, no ocurre lo mismo en $]a, b[$.

2. En **N*** ordenado por $x \mid y$ hay un mínimo elemento 1, pero no hay máximo.

3. $\mathfrak{P}(E)$ ordenado por $A \subset B$ tiene un mínimo elemento $\varnothing$ y un máximo elemento E.

c) Límite superior (resp. inferior) de una parte acotada superiormente (resp. inferiormente) de un conjunto ordenado

Definición. — *Se llama* límite superior en E, *de una parte acotada superiormente* X *de* E *ordenada, la más pequeña de las cotas superiores (si existe)* y límite inferior en E *de una parte acotada inferiormente* X *de* E *la más grande de las cotas inferiores (si existe).*

Estos límites, si existen, son únicos, y se les designa

$$b = \inf_E X \qquad B = \sup_E X,$$

que se enuncia "inf de X en E" y "sup de X en E". Observemos que estas nociones son relativas a X y a E; si

$$X \subset F \subset E$$

X puede tener muy bien un límite superior (o inferior) en E y no tener en F (ver ej. 4 más abajo).

Si no puede surgir ninguna ambigüedad, se escribirá

$$b = \inf X \qquad B = \sup X.$$

EJEMPLOS Y EJERCICIOS

1. Si E es totalmente ordenado y si $a \prec b$

$$\inf \{a, b\} = a \qquad \sup \{a, b\} = b.$$

Igualmente si $X = \{a_1, a_2, \ldots, a_n\}$ con $a_1 \prec a_2 \prec \ldots \prec a_n$ el límite inferior es a_1, y el límite superior es a_n se escribirá

$$\inf (a, b) = a \qquad \inf (a_1, a_2, \ldots, a_n) = a_1$$
$$\sup (a, b) = b \qquad \sup (a_1, a_2, \ldots, a_n) = a_n.$$

2. En **Q** o **R**

$$\inf [a, b] = \inf]a, b[= \inf]a, b] = \inf [a, b[= a$$
$$\sup [a, b] = \sup]a, b[= \sup]a, b] = \sup [a, b[= b.$$

3. En $\mathfrak{P}(E)$ ordenado por $A \subset B$ demostrar que inf $\{X, Y\}$ y sup $\{X, Y\}$ existen y determinarlos.

La misma pregunta para $\{x, y\}$ partes de **N*** ordenada $a \mid b$.

4. Sea X el conjunto de los racionales x positivos tales que $x^2 < 2$, demostrar que X tiene un sup en **R** y no en **Q**.

5. Teniendo X e Y partes de E un límite superior y un límite inferior en E, demostrar que

$$X \subset Y \Rightarrow \inf X \succ \inf Y \quad y \quad \sup X \prec \sup Y.$$

6. Se considera la parte X de **Q**: $(x_n)_{n \in \mathbf{N}^*}$ definida por

$$x_n = \frac{1}{n} + (-1)^n$$

ordenado por $a \leq b$. Demostrar que X tiene un límite inferior y un límite superior.

Con algunos de los ejemplos precedentes se ve que el límite inferior —o superior— de X (si existe) puede o no pertenecer a X. Se demostrará fácilmente el siguiente resultado:

TEOREMA. — *Si* X *es una parte de un conjunto ordenado, las dos propiedades siguientes son equivalentes:*

a) M *es límite superior (resp. inferior) de* X *y pertenece a* X.

b) M *es el máximo (resp. mínimo) elemento de* X.

d) Elementos maximales, elementos minimales de un conjunto ordenado

DEFINICIÓN. — *Si existe un elemento a de* E *ordenado tal que*

$$[x \in E \quad y \quad a \prec x] \Rightarrow x = a$$

se dice que a es un elemento maximal *de* E.

Si existe un elemento a de E *ordenado tal que*

$$[x \in E \quad y \quad x \prec a] \Rightarrow x = a$$

se dice que a es un elemento minimal *de* E.

Dicho de otra manera, un elemento de E ordenado es maximal si no hay en E elementos que le sean estrictamente superiores, y minimal si no hay en E elementos que le sean estrictamente inferiores.

Es evidente que el máximo elemento M de E, si existe, es maximal; por otra parte, es el único; igualmente si hay en E un elemento mínimo m, es minimal y es también el único.

Como se verá en los siguientes ejemplos, las recíprocas son en general falsas; sin embargo, si el conjunto E está totalmente ordenado y si a es un elemento maximal, es comparable todo elemento x de E y es imposible que a admita elementos que le sean estrictamente superiores, y es, pues, superior a todo x de E, luego es el máximo y es único. Análogamente en E totalmente ordenado, si a es un elemento minimal, es el mínimo y es único.

Entonces las nociones de elemento maximal y elemento minimal de un conjunto ordenado no tienen verdaderamente interés más que para los conjuntos parcialmente ordenados.

EJEMPLOS

1. En $\mathscr{P}(E)$ ordenado por $A \subset B$, como hay un elemento mínimo $\varnothing$ y un máximo E, el único elemento minimal es $\varnothing$ y el único elemento maximal es E.

Pero en $\mathscr{P}(E) - \varnothing$ no hay elemento mínimo, las partes $\{x\}$ *con un solo elemento* son los elementos minimales.

En $\mathscr{P}(E) - \{\varnothing, E\}$ (conjunto de las partes propias de E) hay elementos minimales $\{x\}$ y elementos maximales $E - \{x\}$.

2. En $\mathbf{N}^* - \{1\}$ ordenado por $a \mid b$ hay elementos minimales: los números primos no tienen elementos «estrictamente inferiores», es decir, elementos que los dividan y sean diferentes a ellos.

3. En el ejemplo 4 del § 21, la desembocadura del *río* es a la vez el mínimo y el elemento minimal; pero todo nacimiento de *río, de un afluente*, de un *subafluente*, es un elemento maximal.

[illegible]. Orden sobre el producto cartesiano de dos conjuntos ordenados

Dados dos conjuntos ordenados E_1 y E_2 se puede preguntar cómo definir un orden sobre $E_1 \times E_2$ mediante los órdenes definidos sobre E_1 y E_2, se puede hacer de varias maneras, de las que daremos solamente ejemplos.

Ejemplo 1. — Sea E_1 descrito por $x_1, y_1, \dots$ y E_2 descrito por $x_2, y_2, \dots$ Notemos las dos relaciones de orden por la notación $\leq$.

$$x = (x_1, x_2) \qquad y = (y_1, y_2)$$

Consideremos la relación O definida sobre $E_1 \times E_2$ por

$$O(x, y) \leftrightarrow [x_1 \leq y_1 \quad \text{y} \quad x_2 \leq y_2].$$

Se ve fácilmente que la relación O definida sobre $E_1 \times E_2$ es una relación de orden; se dice que este orden es el *orden producto* de los órdenes definidos sobre E_1 y E_2.

Se constata que si E_1 y E_2 tienen cada uno más de un elemento, el orden producto es parcial (si $x_1 < y_1$, $x_2 > y_2$, los pares (x_1, y_2) y (y_1, x_2) no son comparables).

Tomemos, por ejemplo, $E_1 = E_2 = \mathbf{R}$, siendo el orden sobre $\mathbf{R}$ el orden habitual, consideremos el elemento $a = (a_1, a_2)$, las rectas $x_1 = a_1$, $x_2 = a_2$ permiten descomponer el plano $\mathbf{R} \times \mathbf{R}$ en cuatro partes (fig. 7).

I. $x_1 \geq a_1$ y $x_2 \geq a_2$.
II. $x_1 \geq a_1$ y $x_2 < a_2$.
III. $x_1 \leq a_1$ y $x_2 \leq a_2$.
IV. $x_1 < a_1$ y $x_2 > a_2$.

Esta descomposición es un recubrimiento del plan $\mathbf{R} \times \mathbf{R}$, lo que no es una partición, pues I y III tienen por intersección $\{a\}$.

Se observa que $a = (a_1, a_2)$ es comparable con los elementos $x = (x_1, x_2)$ de I (a los que a es inferior) y con los de III (a los que es superior). Pero a no es comparable con los elementos de II y de IV.

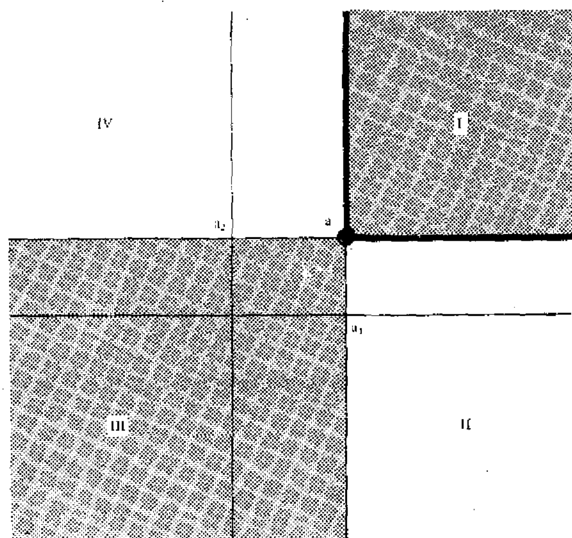


FIG. 7

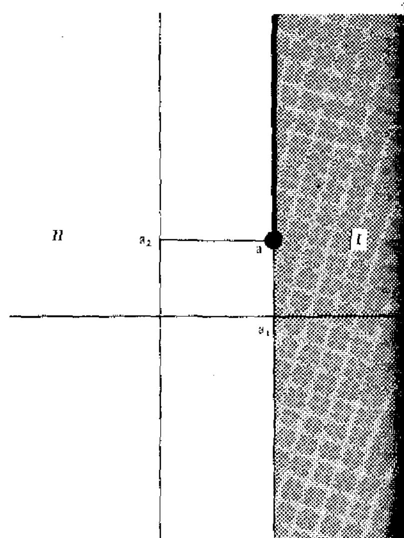


FIG. 8

b) EJEMPLO 2.—Con la misma notación que en el ejemplo precedent *consideremos la relación* O′ definida sobre $E_1 \times E_2$ por

$$O'(x, y) \Leftrightarrow [x_1 < y_1 \quad \text{o} \quad (x_1 = y_1 \quad \text{y} \quad x_2 \leq y_2)].$$

Se prueba fácilmente que O′ es una relación de orden (resulta sólo un po más largo que en el ejemplo precedente). Naturalmente si al menos uno d los órdenes sobre E_1 o E_2 es parcial, el orden definido sobre $E_1 \times E_2$ también parcial; pero si el orden sobre E_1 y E_2 son órdenes totales, se o serva que el orden sobre $E_1 \times E_2$ es también total. Tomemos, por ejempl $E_1 = E_2 = \mathbf{R}$, siendo el orden natural el orden sobre **R**. Consideremos u elemento $a = (a_1, a_2)$, la recta $x_1 = a_1$ y el punto a permiten descompon el plano en dos regiones solamente

$$\text{I.} \quad x_1 > a_1 \qquad \text{o} \qquad (x_1 = a_1 \quad \text{y} \quad x_2 \geq a_2).$$
$$\text{II.} \quad x_1 < a_1 \qquad \text{o} \qquad (x_1 = a_1 \quad \text{y} \quad x_2 \leq a_2).$$

Es un recubrimiento del plano (no es una partición, pues la intersección d I y de II es $\{a\}$). a es comparable a todo punto del plano: es inferior a l puntos de I y superior a los puntos de II.

EJERCICIO

Demostrar utilizando un cambio de coordenadas en el plano que se podría defi en $\mathbf{R} \times \mathbf{R}$ una infinidad de órdenes análogos bien a O, bien a O′.

e) Orden lexicográfico

Podríamos fácilmente extender estos dos ejemplos al producto de más de dos conjuntos ordenados. Consideremos solamente n conjuntos iguales $E_1 = E_2 = \dots = E_n = E$ totalmente ordenados y consideremos el conjunto producto. Pongamos

$$x = (x_1, x_2, \dots, x_n) \qquad y = (y_1, y_2, \dots, y_n)$$

y consideremos la relación L definida por:

— o bien $x_1 < y_1$

— o bien existe un índice $h < n$ tal que

$$x_i = y_i \quad \text{para todo} \quad i < h \quad \text{y} \quad x_h < y_h,$$

— o bien

$$x_i = y_i \quad \text{para todo} \quad i < n \quad \text{y} \quad x_n \leq y_n.$$

[illegible] que la relación L es una relación de orden total, se le llama [illegible], pues siendo E el alfabeto de una lengua ordenada [illegible] (por ejemplo, $a, b, \dots, z$), el orden de las palabras en un [illegible] de esta naturaleza.

[illegible] A en B (A o B ordenados)

[illegible] de un conjunto A en un conjunto B ordenado

[illegible] por una relación que designaremos $\leq$, consideremos [illegible] definida por

$$(\forall x \in A) \qquad f(x) \leq g(x)$$

[illegible] que es una relación de orden, se le designará $f \leq g$.

[illegible] que igualmente si el orden sobre B es total, el orden así [illegible] $\mathcal{F}(A, B)$ es parcial cuando A tiene más de un elemento.

[illegible]ACION

[illegible] significa $f \leq g$ y $f \neq g$, es decir (ver § 11).

$$[(\forall x \in A) \quad f(x) \leq g(x)] \quad \text{y no} \quad [(\forall x \in A) \quad f(x) = g(x)],$$

[illegible]

$$[(\forall x \in A) \quad f(x) \leq g(x)] \quad \text{y} \quad [(\exists x \in A) \quad f(x) \neq g(x)].$$

Por otra parte, si $f(A)$ tiene un límite superior (o inferior) en B, se dirá [illegible] de lenguaje que este límite superior (o inferior) es el límite superior [illegible] de f cuando x describe A y se escribirá

$$\sup_B f(A) = \sup_{x \in A} f \qquad \inf_B f(A) = \inf_{x \in A} f.$$

[illegible]

Igualmente si $f(A)$ es acotada inferiormente (o superiormente, o acotada en B), se dirá, siempre por abuso de lenguaje, que la función f es *acotada inferiormente* (o *acotada superiormente* o *acotada*).

En particular si X es una parte de un conjunto ordenado A, considerando la aplicación canónica de X en A (ver § 12, *b*), se escribirá

$$\sup X = \sup_{x \in X} x \qquad \inf X = \inf_{x \in X} x.$$

EJERCICIO

Demostrar que $\sup_B f(A)$ es el valor de la menor aplicación constante superior a f; igualmente $\inf_B f(A)$ es el valor de la mayor aplicación constante inferior a f.

b) Aplicaciones de un conjunto ordenado A en un conjunto ordenado B

Designaremos con el mismo símbolo las dos relaciones de orden sobre A y B ($\leq$).

DEFINICIÓN 1. — *Se dice que la aplicación f de* A *en* B *es* creciente *si la relación $x_1 \leq x_2$ entraña $f(x_1) \leq f(x_2)$; se dice que f es* decreciente *si la relación $x_1 \leq x_2$ entraña $f(x_1) \geq f(x_2)$.*

Se dice que f es monótona *si f es creciente o si f es decreciente.*

OBSERVACIONES

1. Se observará que si el orden sobre A y B es parcial, la definición de una función decreciente (o creciente) indica que, si x_1 y x_2 son comparables, también lo son $f(x_1)$ y $f(x_2)$.

2. No hay que pensar que una función no creciente es decreciente. En efecto, la negación de «para todo par (x_1, x_2) tal que $x_1 \leq x_2$, se tiene $f_1(x) \leq f_2(x)$» es «existe al menos un par (x_1, x_2) tal que $f(x_1) > f(x_2)$».

DEFINICIÓN 2. — *Se dice que la aplicación f de* A *en* B *es* estrictamente creciente *si la relación $x_1 < x_2$ entraña $f(x_1) < f(x_2)$; se dice que f es* estrictamente decreciente *si la relación $x_1 < x_2$ entraña $f(x_1) > f(x_2)$.*

Se dice que f es estrictamente monótona *si f es estrictamente creciente o estrictamente decreciente.*

TEOREMA. — *Si f y g son monótonas (resp. estrictamente monótonas) $g \circ f$ es monótona (resp. estrictamente monótona); $g \circ f$ es creciente (resp. estrictamente creciente) si y sólo si f y g son las dos crecientes o las dos decrecientes (resp. estrictamente crecientes, o estrictamente decrecientes).*

Pongamos

$$f: \ A \to B; \qquad g: \ B \to C; \qquad h = g \circ f$$
$$x \to y = f(x) \to z = g(y) = h(x).$$

Supongamos, por ejemplo, f creciente y g decreciente, tenemos que

$$x_1 \leq x_2 \Rightarrow y_1 \leq y_2 \Rightarrow z_1 \geq z_2$$

luego h es decreciente.

EJERCICIOS

1. Se considera las aplicaciones siguientes de una parte de **R** en **R**. Determinar los intervalos en que ellas son monótonas

$$x \to f(x) = ax^2 + bx + c \qquad x \to g(x) = \frac{ax+b}{cx+d}$$
$$x \to \log f(x) \qquad x \to e^{f(x)}$$
$$x \to \log g(x) \qquad x \to e^{g(x)}$$
$$x \to \operatorname{arctg} f(x) \qquad x \to \operatorname{arctg} g(x).$$

2. Demostrar que toda aplicación estrictamente monótona de E totalmente ordenado en F ordenado es inyectiva.

c) Isomorfismo por el orden de dos conjuntos ordenados

Dados dos conjuntos ordenados A y B y f una biyección monótona de A sobre B, tenemos

$$x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$$

luego una biyección monótona de A sobre B es estrictamente monótona; por otra parte, siendo f biyectiva y estrictamente monótona, si x_1 y x_2 son comparables, resulta que $f(x_1)$ y $f(x_2)$ son comparables (observación 1 de *b* más arriba), pero si $f(x_1)$ y $f(x_2)$ son comparables, no se puede afirmar en general que x_1 y x_2 son comparables (salvo naturalmente si A es totalmente ordenado), de donde la definición siguiente:

DEFINICIÓN. — *Dados dos conjuntos ordenados* A *y* B, *se llama isomorfismo de* A *sobre* B, *para los órdenes respectivos* A *y* B, *toda biyección tal que las relaciones* $x_1 \leq x_2$ *y* $f(x_1) \leq f(x_2)$ *sean equivalentes.*

Esta definición y las consideraciones anteriores muestran que:

TEOREMA. — *Para que una biyección f de* A *ordenado, sobre* B *ordenado, sea un isomorfismo para los órdenes respectivos definidos sobre* A *y* B, *es condición necesaria y suficiente que f y f^{-1} sean crecientes.*

Los órdenes definidos sobre A *y* B *son los dos parciales, o los dos totales.*

Inversamente, dado un conjunto A ordenado y f una biyección de A sobre B, podemos definir un orden sobre B de la manera siguiente: $y_1 = f(x_1)$, $y_2 = f(x_2)$ de B serán comparables si y sólo si x_1 y x_2 lo son y

$$x_1 \leq x_2 \Leftrightarrow y_1 \leq y_2.$$

Habrá evidentemente un isomorfismo entre A provisto de su orden y B provisto del orden así definido sobre él. Se dice que este orden ha sido definido sobre B por *transporte* del orden definido sobre A.

VII. Conclusión

26. Sobre las nociones de conjunto, de igualdad, ...

Las nociones que hemos definido y los resultados que hemos adquirido relativos a las operaciones sobre los conjuntos, las aplicaciones y las relaciones, nos permiten completar el § 3 mediante las observaciones siguientes:

a) Las dos maneras principales de definir un conjunto

En el § 3, hemos definido un conjunto finito por una enumeración; sea (1) este proceder; y en el § 6, hemos definido un conjunto A contenido en E mediante una propiedad *p* definida sobre E y característica de los elementos de E; sea (2) este *procedimiento*

(1) $$A = \{a, b, c, d, e, f\}$$
(2) $$A = \{x \in E \mid p(x)\}.$$

El procedimiento (1) puede generalizarse así: sean E y F dos conjuntos y *f* una aplicación de E en F, se puede deducir de ella (ver § 16) una suprayección *s* de E sobre $A = f(E) = s(E)$

$$E \to A = s(E)$$
$$t \to x = s(t).$$

Se dice que *s* es una *representación paramétrica* de A mediante E, llamado *conjunto de parámetros* (procedimiento 1′).

EJEMPLOS

1. Consideremos la aplicación *f* de **R** en $\mathbf{R}^2$ definida por

$$t \to (x, y) \qquad x = \cos t \qquad y = \operatorname{sen} t.$$

Si los elementos de $\mathbf{R}^2$ se expresan en función de un sistema de referencia ortonormal, $f(\mathbf{R})$ será el conjunto de los puntos de un círculo de centro (0, 0) y de radio 1.

2. Consideremos la aplicación F de $\mathbf{R}^2$ en $\mathbf{R}^3$ definida por

$$(u, v) \mapsto (x, y, z) \qquad x = f(u, v) \qquad y = g(u, v) \qquad z = h(u, v)$$

siendo *f, g, h* tres funciones continuas de las dos variables reales *u* y *v*.

Se verá más adelante que con ciertas condiciones $F(\mathbf{R}^2)$ está constituido por un conjunto de puntos llamado *superficie*; por ejemplo, en $\mathbf{R}^3$ expresado en una referencia ortonormal si

$$\begin{cases} x = \cos u \cos v \\ y = \cos u \operatorname{sen} v \\ z = \operatorname{sen} u \end{cases}$$

$F(\mathbf{R}^2)$ es una esfera de centro (0, 0, 0) y de radio 1.

Este procedimiento (1′) de definir A mediante una suprayección s: E $\to$ A se parece a (1); en efecto, conociendo los elementos de E, x, y, ... conocemos los elementos $s(x)$, $s(y)$, ... de A por así decir uno por uno como en la enumeración de un conjunto finito. Los métodos (1) (1′) nos "proporcionan" los elementos de A, pero no nos indican de estos elementos más propiedad que la siguiente: $x \in$ A. Si queremos definir otras propiedades de los elementos de A, será preciso demostrar *teoremas*.

Consideremos, por el contrario, un conjunto definido por el procedimiento 2, conocemos una propiedad característica de los elementos de A, pero no conocemos ningún elemento de A, ni siquiera sabemos si A es vacío o no, se nos presenta un *problema* para resolver: encontrar los elementos x de E tales que x posea p. Si la resolución de este problema nos da A $\neq \varnothing$, enunciaremos un *teorema de existencia*. En el caso particular en que A $= \{x\}$, la solución del problema precedente nos conducirá a enunciar un *teorema de unicidad*. Los teoremas de existencia y de unicidad son seguidos a menudo de una definición.

EJEMPLOS

1. Consideremos $\mathbf{R}^2$ expresado mediante una referencia ortonormal, sea A el conjunto de los (x, y) tales que

$$x^2 + y^2 = 1.$$

Si conocemos las funciones circulares vemos que

$$\mathrm{A} = \{(x, y) \mid x^2 + y^2 = 1\} = \{(x, y) \mid (\exists t \in \mathbf{R}) \quad x = \cos t \quad y = \operatorname{sen} t\}.$$

Hemos mostrado la existencia de los elementos de A, y demostrado, además, la equivalencia para este ejemplo del procedimiento (1′) y del procedimiento (2) de definición de A.

Con frecuencia el enunciado del problema debe ser precisado por una discusión antes de enunciar un teorema de existencia o de unicidad.

2. Siendo a y b dos reales encontrar la parte A de $\mathbf{R}$ descrita por x tal que $ax + b = 0$

$a \neq 0$	A $= \{-b/a\}$
$a = 0$ y $b \neq 0$	A $= \varnothing$
$a = b = 0$	A $= \mathbf{R}$.

b) Conjuntos deducidos de un conjunto E

Dado E, podemos definir las partes de E: A, B, ... después el conjunto $\mathfrak{P}$(E), los conjuntos A $\cup$ B, A $\cap$ B, conjuntos productos A $\times$ B, conjuntos cocientes E/R, conjuntos de aplicaciones $\mathfrak{F}$(A, B).

Estando definidos todos estos conjuntos, podemos aplicarles estos mismos procedimientos y recomenzar indefinidamente.

En nuestro nivel, todos los conjuntos que vamos a considerar serán deducidos así de $\mathbf{N}$: por ejemplo, $\mathbf{Z}$ será $(\mathbf{N} \times \mathbf{N})/\mathbf{R}$, $\mathbf{Q}$ será $(\mathbf{Z} \times \mathbf{Z}^*)/\mathbf{R}'$ (siendo $\mathbf{R}$ y $\mathbf{R}'$ las relaciones de equivalencia convenientemente escogidas (ver § 58

y § 107)). **R** será descrito por los elementos de $\mathfrak{F}$(**Q**), **C** será **R** × **R**, estos conjuntos **R** y **C** están provistos de ciertas operaciones (ver § *109* y § *113*). El conjunto de los puntos del espacio de la geometría elemental será $\mathbf{R}^3$, ...

c) Noción de igualdad

Nosotros hemos dicho en el § 3 que dos objetos son iguales si son idénticos. Por ejemplo, un elemento que está representado por un símbolo más o menos complicado se le puede representar por un símbolo más simple, así si $f(x)$ es la imagen de x por f, se escribirá $y = f(x)$, los dos objetos y y $f(x)$ son idénticos.

Hemos visto que una propiedad de un elemento x definido sobre un conjunto puede ser verdadera o falsa según la elección de x en E; sucede a menudo que esta propiedad se expresa mediante una igualdad, entonces tenemos una ecuación (ver § 14).

$$f(x) = b.$$

Se dice que es una *igualdad condicional,* se transforma en una igualdad si $x \in f^{-1}(b)$, es decir, si x es solución de la ecuación.

En una teoría, dos elementos distintos de un conjunto E pueden hacer un papel análogo; hemos sustituido esta noción vaga por la de *equivalencia módulo* R (R relación de equivalencia definida sobre E)

$$a \equiv b \qquad (\text{mod R}) \Leftrightarrow \dot{a} = \dot{b}.$$

Sucede a veces que se confunde una clase de equivalencia y un representante, es un abuso de lenguaje frecuente que es preciso evitar; pero si se comete hay que ser consciente de ello.

EJEMPLOS

1. Cuando escribimos (ver § 18, ej. 6)

(1) $$\overrightarrow{AB} = \overrightarrow{CD}$$

no queremos, en general, decir que los vectores ligados $\overrightarrow{AB}$ y $\overrightarrow{CD}$ son idénticos, sino que son equivalentes; es decir, el «vector libre representado por $\overrightarrow{AB}$ es idéntico al vector libre representado por $\overrightarrow{CD}$»; haría falta escribir

(1′) $$\left(\dot{\overrightarrow{AB}}\right) = \left(\dot{\overrightarrow{CD}}\right)$$

pero los vectores ligados se usan poco, se escribe «la igualdad» convencional (1) por la (1′).

2. Veremos que para las fracciones a/a', b/b', la relación $ab' = a'b$ es una relación de equivalencia, la clase de equivalencia $(\dot{a/a'})$ es un número racional cuando escribimos

(2) $$\frac{9}{6} = \frac{3}{2}.$$

No queremos decir con esto que las fracciones 9/6 y 3/2 son idénticas (lo que es falso), sino que

$$(2') \qquad \left(\dot{\frac{9}{6}}\right) = \left(\dot{\frac{3}{2}}\right).$$

es decir, que son dos representantes de un mismo número racional.

d) Observación sobre el verbo ser

El uso del verbo ser es un excelente ejemplo de la precisión obtenida al pasar del lenguaje corriente al lenguaje matemático. Consideremos los enunciados siguientes:

1. 2 *es* el número de los elementos de $\{a, b\}$ $a \neq b$ (le diremos el *cardinal* de $\{a, b\}$, ver § 31).
2. a *es* un número par.
3. Los números pares *son* enteros.

Si P designa el conjunto de los enteros pares, estos enunciados traducen:

1. Una igualdad	$2 = \text{card}\ \{a, b\}$ $(a \neq b)$.
2. Una pertenencia	$a \in \mathbf{P}$.
3. Una inclusión	$\mathbf{P} \subset \mathbf{Z}$.

La confusión de estas tres nociones es la causa de las paradojas surgidas en la lógica tradicional.

En conclusión, preferiremos siempre al lenguaje corriente, forzosamente ambiguo, la precisión del lenguaje matemático y si empleamos abusos de lenguaje "sin los que todo texto matemático resultaría pedante e incluso ilegible" (N. BOURBAKI), los señalaremos como hemos hecho en este primer capítulo.

Ejercicios

1. Dados tres conjuntos A, B, C *tales* que

$$A \cup B \subset A \cup C \qquad \text{y} \qquad A \cap B \subset A \cap C$$

¿qué se puede decir de los dos conjuntos B y C?

2. Dadas dos partes A y B de E, se llama *diferencia simétrica* de A y B (que se escribirá $A \Delta B$) el conjunto

$$A \Delta B = (A \cup B) - (A \cap B).$$

a) Demostrar que $A \Delta B = (A - B) \cup (B - A)$.

b) Demostrar que, cualesquiera que sean tres partes A, B, C de E,

$$(A \Delta B) \Delta C = A \Delta (B \Delta C).$$

c) Demostrar que existe una sola parte X de E tal que, 1) para toda parte A de E: $A \Delta X = X \Delta A = A$, 2) existe una sola parte A′ de E tal que $A \Delta A' = A' \Delta A = X$.

3. En el plano de la *geometría* elemental se suponen conocidas las definiciones de una traslación, de una rotación, de un desplazamiento, de una *semejanza*. Se designa, *respectivamente*, por $\mathfrak{T}$, $\mathfrak{R}$, $\mathfrak{D}$, $\mathfrak{H}$, $\mathfrak{S}$ el conjunto de las traslaciones, de las rotaciones, de los desplazamientos, de las homotecias, de las semejanzas del plano. Determinar las intersecciones dos a dos de estos cinco conjuntos.

4. Siendo f una aplicación de A en B (A y B no vacíos), demostrar que las dos propiedades siguientes son equivalentes: **a)** f es inyectiva; **b)** existe f' de **B** *en* **A** tal que $f' \circ f = \mathrm{id}_A$.
Cuando f' existe, ¿es única? Estudiar el caso de $A = B = \mathbf{N}$ con $f(n) = 2n$.

5. Siendo g una aplicación de A en B (A y B no vacíos), demostrar que las dos propiedades siguientes son equivalentes: **a)** g es suprayectiva; **b)** existe g' de B en A tal que $g \circ g' = \mathrm{id}_B$.
Si g' existe, ¿es única? Demostrar que, si existe g'_1 y g'_2 tales que $g'_1(B) = g'_2(B)$, entonces $g'_1 = g'_2$. Estudiar el caso $A = B = \mathbf{N}$ con $g(2p) = p$, $g(2p + 1) = 0$.

6. Sean tres aplicaciones $f : \mathbf{A} \to \mathbf{B}$, $g : \mathbf{B} \to \mathbf{C}$, $h : \mathbf{C} \to \mathbf{A}$, demostrar que si entre las tres aplicaciones $h \circ g \circ f$, $g \circ f \circ h$, $f \circ h \circ g$ dos son inyectivas (resp. suprayectivas), y la tercera suprayectiva (resp. inyectiva), entonces f, g, h, son biyectivas.

7. Sean tres aplicaciones $f : \mathbf{A} \to \mathbf{B}$, $g : \mathbf{B} \to \mathbf{C}$, $h : \mathbf{C} \to \mathbf{D}$ tales que $g \circ f$ y $h \circ g$ sean biyectivas, demostrar que f, g, h son biyectivas.

8. Siendo a un entero natural estrictamente positivo, se consideran las aplicaciones de $\mathbf{N}^*$ en sí mismo definidas por

$$y = f(x) = D(a, x) \qquad z = g(x) = M(a, x).$$

Donde $D(a, x)$ y $M(a, x)$ designan el M.C.D. y el M.C.M. de a y x, respectivamente. Determinar $f(\mathbf{N}^*)$, $g(\mathbf{N}^*)$, $f^{-1}(y)$ y $g^{-1}(z)$.

9. Siendo A una parte fija del conjunto E, se consideran las aplicaciones de $\mathfrak{P}(E)$ en sí mismo definidas por

$$Y = f(X) = A \cap X \qquad Z = g(X) = A \cup X$$

Determinar

$$f(\mathfrak{P}(E)), \qquad g(\mathfrak{P}(E)), \qquad f^{-1}(Y) \qquad \text{y} \qquad g^{-1}(Z).$$

10. Sean **A** y **B** dos conjuntos no vacíos.

a) Siendo $\mathbf{X}_1$ y $\mathbf{X}_2$ dos partes *no vacías* de **A** y f_1 y f_2 dos aplicaciones $f_1 : \mathbf{X}_1 \mapsto \mathbf{B}$, $f_2 : \mathbf{X}_2 \to \mathbf{B}$ demostrar que, para que exista una aplicación f de $\mathbf{X}_1 \cup \mathbf{X}_2$ en **B** tal que f_1 sea la restricción de f a $\mathbf{X}_1$ y f_2 la restricción de f a $\mathbf{X}_2$, es *necesario* y *suficiente* que para todo x de $\mathbf{X}_1 \cap \mathbf{X}_2$, $f_1(x) = f_2(x)$.

b) Generalizar a una familia de partes $(\mathbf{X}_i)$ de **A** y a una familia de aplicaciones (f_i) de $\mathbf{X}_i$ en **B**, $(\mathbf{X}_i)$ y (f_i) referidos al mismo conjunto de índices **I**. Demostrar que para que exista f

$$f\colon \left(\bigcup_{i \in \mathbf{I}} \mathbf{X}_i\right) \to \mathbf{B}$$

tal que para todo i de **I**, f_i sea la restricción de f a $\mathbf{X}_i$, es necesario y suficiente que para todo par $(i, j) \neq (i, j)$ y para todo x de $\mathbf{X}_i \cap \mathbf{X}_j$, $f_i(x) = f_j(x)$.

11. Con las mismas notaciones que en el ejercicio 9, determinar las relaciones de equivalencia R y S asociadas, respectivamente, a las aplicaciones f y g; determinar $\mathfrak{F}(E)/R$ y [illegible] y demostrar que existe una biyección de $\mathfrak{F}(E)/\mathbf{R}$ sobre $\mathfrak{F}(A)$ y una biyección de [illegible] sobre $\mathfrak{F}(E - A)$.

12. [illegible] el error en el razonamiento siguiente: «R es una relación binaria definida [illegible] y transitiva; luego $\mathbf{R}(x, y) \to \mathbf{R}(y, x)$ (simetría) y $\mathbf{R}(y, x) \Rightarrow \mathbf{R}(x, x)$ [illegible] luego R es reflexiva y, por consiguiente, es una relación de equi-[illegible]

[illegible] de E en F (E y F no vacíos), existe una relación de equiva-[illegible] se define sobre E la relación R por

$$\mathbf{R}(x, y) \Leftrightarrow \mathbf{S}(f(x), f(y)).$$

[illegible] R es una relación de equivalencia. ¿Cuáles son sus clases? ¿Es com-[illegible]

[illegible] conjuntos E y F no vacíos provistos, respectivamente, de una relación de [illegible] R, y de una relación de equivalencia S, se designa por u y v las aplica-[illegible] u: $E \to E/R$, v: $F \to F/S$. Se dirá que f, aplicación de E en F, es [illegible] con R y S si $v \circ f$ es compatible con R. Demostrar que en este caso: [illegible] x' (mod R) $\to f(x) \equiv f(x')$ (mod S).
[illegible] h: $E/R \to F/S$ tal que $v \circ f = h \circ u$.
[illegible]

[illegible] R una relación de equivalencia definida sobre E, se dice que una parte S [illegible] está *saturada* por R si, para todo x de S, S contiene la clase de x módulo R. [illegible] *parte saturada engendrada por una parte* A de E y se designa sat (A), la más pequeña parte saturada de E conteniendo A.

a) Demostrar que sat $(A) = \bigcup_{x \in A} \dot{x}$

b) Siendo R la relación de equivalencia *asociada* a una aplicación f de E en un conjunto cualquiera F, demostrar que, para que S esté saturado por R, es necesario y suficiente que $f^{-1}(f(S)) = S$.
Siendo A una parte de E, determinar sat (A) para este relación **R**. Demostrar que hay una biyección entre las partes saturadas de E y las partes de $f(E)$.

[illegible] Dada una aplicación f de E en F, se designa por $\mathcal{S}$ la familia de partes de S de E tales que $f^{-1}(f(S)) = S$.

a) Siendo A una parte cualquiera de E, demostrar que $f^{-1}(f(A))$ es un conjunto de $\mathcal{S}$.

b) Demostrar que toda intersección y toda reunión de conjuntos de $\mathcal{S}$ es un conjunto de $\mathcal{S}$.

c) Siendo S un conjunto de $\mathcal{S}$ y A una parte de E tal que S y A sean disjuntos, demostrar que $\mathcal{S}$ y $f^{-1}(f(A))$ son disjuntos.

d) Siendo S_1 y S_2 dos conjuntos de $\mathcal{S}$ tales que $S_1 \subset S_2$, demostrar que $S_2 - S_1$ es un conjunto de $\mathcal{S}$.

17. Siendo R_1 una relación de equivalencia definida sobre un conjunto E_1 y R_2 una relación de equivalencia definida sobre un conjunto E_2, se designa por **R** la relación entre

$x = (x_1, x_2)$ e $y = (y_1, y_2)$ de $E = E_1 \times E_2$ definido por $R_1(x_1, y_1)$ y $R_2(x_2, y_2)$.

a) Demostrar que **R** es una relación de equivalencia; ella se llama equivalencia producto de $\mathbf{R}_1$ y $\mathbf{R}_2$ y designada $\mathbf{R}_1 \times \mathbf{R}_2$.

b) Demostrar que: $(E_1 \times E_2)/(\mathbf{R}_1 \times \mathbf{R}_2) = (E_1/\mathbf{R}_1) \times (E_2/\mathbf{R}_2)$.

18. Dado un conjunto E, se designa por $\mathcal{R}$ el conjunto de las relaciones de equivalencia definidas sobre E; para simplificar se representará con la misma letra una equivalencia y su grafo.

a) *Siendo* $\mathbf{R}_1$ *y* $\mathbf{R}_2$ *dos elementos de* $\mathcal{R}$, se considera la relación **R**: «$\mathbf{R}_1$ y $\mathbf{R}_2$» llamada intersección de $\mathbf{R}_1$ y $\mathbf{R}_2$. Demostrar que **R** es una relación de equivalencia. ¿Cuál es su grafo? Demostrar que una clase módulo **R** es la intersección de una clase módulo $\mathbf{R}_1$ y de una clase módulo $\mathbf{R}_2$.
Estudiar los ejemplos siguientes: α) $E = \mathbf{Z}$, $\mathbf{R}_1$ y $\mathbf{R}_2$ siendo las congruencias de módulos respectivos p_1 y p_2 (p_1 y p_2 enteros estrictamente positivos distintos). β) E es el plano de la geometría elemental. R_1 y R_2 son, respectivamente, las relaciones «MM′ es paralela a d_1», «MM′ es paralela a d_2» (d_1, d_2 son dos direcciones distintas del plano).

b) Con las misma notaciones anteriores, averiguar si la relación «$\mathbf{R}_1$ o $\mathbf{R}_2$» es una relación de equivalencia. ¿Cuál es su grafo?

c) Siendo $(\mathbf{R}_i)$ una familia de relaciones de equivalencias definidas sobre E, cuyo conjunto de índices es I, se designa por R la relación siguiente, llamada *intersección* de las $(\mathbf{R}_i)$,

$$\mathbf{R}(x, y) \Leftrightarrow \text{para todo } i \text{ de I, } \mathbf{R}_i(x, y).$$

Demostrar que **R** es una relación de equivalencia. ¿Cuál es su grafo? Demostrar que una clase C módulo R está contenida para todo i en una y sólo una clase. C_i módulo R_i y que C es la intersección de la familia (C_i).

19. Tomamos las mismas notaciones generales que en el ejercicio 18. Se dice que **R** es *«más fina»* que **R′** si y solamente si

$$\mathbf{R}(x, y) \Rightarrow \mathbf{R}'(x, y).$$

a) Demostrar que esta relación entre **R** y **R′** define un orden sobre $\mathcal{R}$; comparar los grafos de R y R′; este orden, ¿es total o parcial?

b) Demostrar que **R** es más fina que **R′** si y solamente si toda clase de equivalencia módulo **R′** es un reunión de clases de equivalencia módulo **R**.

c) ¿Existe un elemento mínimo y uno máximo en $\mathcal{R}$ ordenado por la relación de orden definido más arriba?

d) E = **Z**, determinar todas las congruencias módulo un entero estrictamente positivo que son más finas que $x \equiv y$ (módulo n) o que son menos finas que esta relación.

20*. Con las mismas notaciones que en los ejercicios 18 y 19.

a) Demostrar que el conjunto de las relaciones de equivalencias más finas que $\mathbf{R}_1$ y $\mathbf{R}_2$ no es vacío y que entre estas últimas hay una menos fina que las otras, a saber, «$\mathbf{R}_1$ y $\mathbf{R}_2$» que es, pues, inf $(\mathbf{R}_1, \mathbf{R}_2)$.

b) Demostrar que el conjunto de las relaciones de equivalencia menos finas que R_1 y R_2 no es vacío y su intersección S (ver ej. 18 **c**)) no es vacía. Demostrar que esta relación **S** es la más fina de todas las que son menos finas que $\mathbf{R}_1$ y $\mathbf{R}_2$, es decir, **S** es sup $(\mathbf{R}_1, \mathbf{R}_2)$. Determinar la relación **S** para los dos ejemplos α y β (ej. 18, **a**).

c) *Se considera la relación* **S'** *definida sobre* E *de la manera siguiente:* «$\mathbf{S}'(x, y)$ si y solamente si existe una sucesión $a_1, a_2, \ldots, a_n$ de elementos de E tales que

$$[\mathbf{R}_1(x, a_1) \ \text{ o } \ \mathbf{R}_2(x, a_2)] \ \text{ y } \ [\mathbf{R}_1(a_1, a_2) \ \text{ o } \ \mathbf{R}_2(a_1, a_2)] \ \text{ y } \ldots$$

y $[\mathbf{R}_1(a_n, y)$ o $\mathbf{R}_2(a_n, y)]$». Demostrar que **S'** es una relación de equivalencia definida sobre E *a la vez menos fina y más fina que la relación* **S** definida en la pregunta **b**) anterior, luego que S' es equivalente a S = sup (R_1, R_2).

21. Siendo A y B dos conjuntos, se designa por F el conjunto de las aplicaciones de una parte de A en B, siendo f un elemento de F, se designará por D(f) el conjunto de definición de f.

a) Demostrar que la relación definida sobre F, «g prolonga f», es decir, «$D(g) \supset D(f)$ y la restricción de g a D(f) es f» es una relación de orden sobre F. Este orden, ¿es total o parcial.

b) Determinar los elementos máximos de F para el orden así definido.

c) Siendo f y g dos elementos de F (o bien (f_i) una familia de elementos de F con I como conjunto de índices), determinar en qué condiciones existe sup (f, g) o $\sup_i (f_i)$. (Utilizar el ejercicio 10.)

22. Dada la aplicación f de A en B, la relación $y = f(x)$ es compatible respecto a x con la relación de equivalencia $\mathbf{R} : f(x) = f(x')$. Demostrar que hay otras relaciones de equivalencia definidas sobre A que poseen esta propiedad; caracterizar **R** entre estas últimas con la noción de finura (ver ejercicio 19).

23. Se dice que una relación binaria definida sobre E es una relación de *preorden* si es reflexiva y transitiva; sea **P** una relación tal. Se considera la relación **R** definida sobre E por

$$\mathbf{R}(x, y) \Leftrightarrow \mathbf{P}(x, y) \text{ y } \mathbf{P}(y, x).$$

a) Demostrar que **R** es una relación de equivalencia.

b) Demostrar que **P** es compatible con **R**, deducir que se puede definir sobre E/**R** una relación **O** mediante

$$\mathbf{O}(\dot{x}, \dot{y}) \Leftrightarrow \mathbf{P}(x, y).$$

Demostrar que **O** es una relación de orden definida sobre E/**R**. Se la llama relación de orden **O** asociada a la relación **P**.

c) Demostrar que las relaciones siguientes son relaciones de preorden, buscar las relaciones de orden asociadas.

a) E = **R** × **R** la relación P entre $x = (x_1, x_2)$ e $y = (y_1, y_2)$ siendo $x_1 \leq y_1$.

β) $E = \mathfrak{F}$, $\mathfrak{F}$ conjunto de partes finitas de **N**, se designa por $\delta(X)$ la suma de los enteros de una parte finita X de **N**, la relación P es $\delta(X) \leq \delta(Y)$.
γ) $E = \mathbf{Z}^*$ la relación **P** es «x divide y».

24. Se dice que un conjunto ordenado T es un *retículo* (o un *conjunto reticulado*) si y sólo si existe para todo par (x, y) de elementos de T, inf (x, y) y sup (x, y). Todo conjunto totalmente ordenado es un retículo, luego la noción de retículo es interesante principalmente para los conjuntos parcialmente ordenados.
Demostrar que cada uno de los conjuntos siguientes parcialmente ordenados es un retículo, determinar en cada caso inf (x, y) y sup (x, y).

a) $\mathfrak{P}(E)$ ordenado por $A \subset B$.

b) $\mathbf{N}^*$ ordenado por $x \mid y$ (x divide y).

c) $\mathcal{R}$ conjunto de las relaciones de equivalencia ordenadas por «R es más fino que R′» (ver ejercicios 18 y 19).

d) $\mathfrak{L}$ conjunto de las partes del espacio E de la geometría elemental comprendiendo únicamente: la parte vacía y la parte llena de E, las partes de E reducidas a un punto, las rectas y los planos. La relación de orden es la inclusión inducida sobre $\mathfrak{L}$ por la inclusión definida sobre $\mathfrak{P}(E)$.

25. Sea (a_i, b_i) una familia finita $(1 \leq i \leq n)$ de intervalos (cerrados, abiertos o semiabiertos), de **R** tales que dos cualesquiera de entre ellos se cortan. Demostrar que la intersección **I** de dicha familia es un intervalo cerrado no vacío (cerrado, abierto o semiabierto) que se determinará. ¿Se podría en este enunciado reemplazas **R** por un conjunto cualquiera E totalmente ordenado?

2

ENTEROS NATURALES

I. Conjunto **N** de los enteros naturales.
II. Conjuntos finitos.
III. Operaciones sobre los enteros naturales.
IV. Análisis combinatorio.
V. Nociones sobre los conjuntos numerables.

I. Conjunto N de los enteros naturales

27. Los programas de estudios (Ciencias y escuelas especiales) correspondientes al nivel intelectual a que este libro va dirigido autorizan al autor a admitir la existencia del conjunto de los números enteros **N**, sus propiedades y las operaciones clásicas entre sus elementos.

Sin embargo, todos los demás conjuntos de números estudiados en este curso **Z**, **Q**, **R**, **C** al construirse a partir de **N**, nos ha parecido que un cierto número de estudiantes estarían interesados en la exposición de un modo de introducción de **N**.

Por otra parte, los que no sigan sus estudios matemáticos hasta el final de la licenciatura y se dediquen a la enseñanza de la Aritmética les agradará conocer uno de los sistemas de axiomas que definen **N** y sus consecuencias: no para que enseñen estos axiomas a sus jóvenes alumnos, sino para que sepan *algo más* que ellos en esta *materia*.

La existencia de **N** deriva de la teoría de conjuntos. Para simplificar *admitimos la existencia de* **N**, contentándonos con determinar las propiedades clásicas partiendo de un reducido número de ellas (sección I). Estudiaremos seguidamente los conjuntos (sección II). Esto nos permitirá definir correctamente las operaciones en **N** (sección III).

A continuación con los resultados así adquiridos y con la noción de aplicación, demostraremos un cierto número de propiedades de los conjuntos finitos conocidas bajo el nombre de *análisis combinatorio*. Esta sección reitera el estudio de resultados en parte ya conocidos por el lector, si bien se presentan a partir de la teoría de conjuntos.

En fin, en la sección V, daremos algunas indicaciones sobre los conjuntos numerables.

28. Introducción de los enteros naturales

Designemos por E un conjunto no vacío poseyendo las tres propiedades designadas más abajo por N_1, N_2, N_3.

Axioma N_1. — E *es un conjunto ordenado, toda parte no vacía de* E *tiene un elemento mínimo.*

Consecuencias:

1. Toda parte $\{x, y\}$ de E tiene un elemento mínimo, luego E *es totalmente ordenado.*
2. La parte llena de E tiene un elemento mínimo, luego E *tiene un elemento mínimo* m.

Axioma N_2. — E *no tiene elemento máximo.*

Consecuencia: Todo elemento a *de* E *tiene un sucesor* a' (ver § 22, *c*).

Sea X el conjunto de los elementos x de E inferiores o iguales a a e Y el conjunto de las cotas superiores de X no pertenecientes a X, es el complementario de X con respecto a E, Y no es vacío, si no a sería el máximo elemento de E; luego Y tiene un elemento mínimo (axioma N_1), sea a'; el intervalo abierto $]a, a'[$ es vacío, pues si existiera b tal que $a < b < a'$, b sería una cota superior de X que no pertenecería a X, luego un elemento de Y y a' no sería el mínimo elemento de Y.

Axioma N_3. — *Todo elemento* a *de* E *distinto de* m *tiene un predecesor* $'a$ (ver § 22, *c*).

Consecuencia: Toda parte no vacía de E *acotada superiormente tiene un elemento máximo.*

Sea X una parte no vacía acotada superiormente de E, y sea Z el conjunto de las cotas superiores de X, Z no es vacío, puesto que X es acotado superiormente, y tiene, por lo tanto, un elemento mínimo a (que es la menor de las cotas superiores de X, es decir, sup X).

Si $a = m$, $X = \{m\}$, m es entonces también el mayor (y único elemento de X).

Si $a \neq m$, a pertenece a X; en efecto, si a no perteneciera a X, su predecesor $'a$ sería también cota superior de X, puesto que $]'a, a[$ es vacío, y a no sería la más pequeña de las cotas superiores de X.

En resumen, un conjunto E verificando los axiomas N_1, N_2, N_3 tiene las propiedades siguientes:

— E *es totalmente ordenado.*

— E *tiene un elemento mínimo* m *y no tiene elemento máximo.*

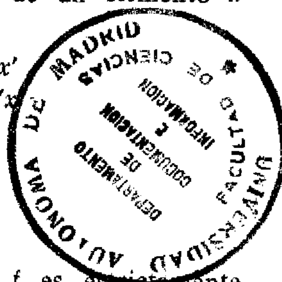
— *Todo elemento* x *de* E *tiene un sucesor* x' *y todo elemento* $x \neq m$ *de* E *tiene un predecesor* $'x$.

—*Toda parte no vacía de* E *tiene un elemento mínimo y toda parte acotada superiormente no vacía de* E *tiene un elemento máximo.*

Para simplificar la exposición (ver ejercicio 28 fin del capítulo), *admitimos que todos los conjuntos* E *que poseen las propiedades* N_1, N_2, N_3 *son isomorfos para la estructura de orden y designamos por* N *aquel cuyo primer elemento se representa por* 0, *el sucesor de* 0 *es* 1, *el de* 1 *es* 2, *etc.* Los elementos de N se llaman *enteros naturales.* Se representa $N^* = N - \{0\}$.

Observemos también que el sucesor y el predecesor de un elemento x siendo únicos (ver § 22, *c*), las aplica iones

$$f: N \to N^* \quad \text{definido por} \quad f(x) = x'$$
$$g: N^* \to N \quad \text{definido por} \quad g(x) = {}'x$$



son biyecciones recíprocas una de otra, pues

$$(\forall x \in N^*) \qquad ('x)' = x$$
$$(\forall x \in N) \qquad '(x') = x.$$

Por otra parte, $x <$ _ implica $x < x' \leq$ _ $< y'$, luego f es estrictamente creciente; análogamente con g.

Toda familia de elementos (ver § 17) cuyo conjunto de índices es N o una parte de N se llama una *sucesión.* Una sucesión $a_n (n \in N)$ es *estacionaria* si existe p tal que $n > p$ entraña $a_n = a_p$.

Todo conjunto equipotente (ver § 20) con N o con una parte de N se llama *numerable* (ver § 42).

19. Noción de recurrencia

Ocurre que se toma por uno de los axiomas definiendo N el resultado siguiente, que es el principio de razonamiento por recurrencia; con el modo de exposición adoptado, este enunciado se transforma en un teorema:

Teorema.— *Toda parte* X *de* N *tal que*

$$0 \in X \qquad y \qquad (x \in X \Rightarrow x' \in X)$$

es idéntico a N.

Sea Y el complementario de X con respecto a N, si Y es vacío, el teorema está demostrado; supongamos, pues, Y no vacío, Y tiene, por lo tanto, un elemento mínimo a (axioma N_1), a es distinto de 0, puesto que 0 pertenece a X, luego $'a$ existe y es estrictamente inferior a todo elemento de Y, luego no pertenece a Y, sino a X; en virtud de la segunda hipótesis $('a)' = a$ pertenece a X, lo que es imposible, puesto que a pertenece a Y, luego la hipótesis hecha sobre Y es falsa: Y es, pues, vacío y $X = N$.

Hemos enunciado este resultado utilizando las nociones de elementos y de partes del conjunto N, podemos también enunciarlo mediante la noción

de propiedad definida sobre N (ver § 6). Sea, en efecto, p una propiedad definida sobre N escribiendo

$$X = \{x \in N \mid p(x)\}.$$

Obtenemos el teorema siguiente, fundamento del razonamiento por recurrencia:

TEOREMA. — *Si p es una propiedad definida sobre* N *tal que*

$$p(0) \text{ y } [(\forall x \in N) \ (p(x) \Longrightarrow p(x'))]$$

entonces p es verdadero para todo x de N.

Este teorema sirve también para construir una sucesión de elementos de un conjunto A en las condiciones siguientes: se conoce a_0 y conociendo a_n, se posee un método que permite formar $a_{n'}$. Designemos por p la propiedad siguiente: "se sabe formar a_n de A"; según las condiciones dadas, p verifica las hipótesis del teorema precedente, luego p es cierta para todos los enteros naturales y sabemos construir, en consecuencia, la sucesión (a_n) $(n \in N)$.

En los teoremas precedentes, las hipótesis sirven para mostrar que $X = N$ o que p es verdadera para todos los enteros naturales, se dice entonces que se efectúa una recurrencia sobre N; con hipótesis distintas se obtendrá los resultados siguientes que el lector demostrará fácilmente:

COROLARIO 1. — *Si a es un entero natural distinto de 0, toda parte* X *de* N *tal que*

$$a \in X \quad y \quad [(\forall x \in X) \ x' \in X]$$

contiene $[a, \rightarrow[$ (recurrencia a partir de a).

COROLARIO 2. — *Si b es un entero distinto de 0, toda parte de* N *tal que*

$$0 \in X \quad y \quad [(\forall x \in [0, b[) \ x' \in X]$$

contiene $[0, b]$ (recurrencia hasta b).

COROLARIO 3. — *Si a y b son dos enteros naturales tales que $0 < a < b$, toda parte* X *de* N *tal que*

$$a \in X \quad y \quad [(\forall x \in [a, b[) \ x' \in X]$$

contiene $[a, b]$ (recurrencia sobre un intervalo).

II. Conjuntos finitos

30. Definición. Partes finitas de N

DEFINICIÓN. — *Un conjunto* F *es* finito, *si es vacío o si existe una biyección de* F *sobre un intervalo cerrado $[1, a]$ de* N *$(a \neq 0)$. Un conjunto no finito se llama* infinito.

Estudiaremos primero en este párrafo las *partes finitas* de N.

Teorema 1. — *Si a y b son dos enteros naturales no nulos, si existe una inyección de* $A = [1, a]$ *en* $B = [1, b]$, $a \leq b$.

Razonemos por recurrencia sobre a, el teorema es verdadero para $a = 1$, puesto que $b \geq 1$.

Supongamos el teorema cierto para a, siendo a' el sucesor de a; consideremos una inyección f de $A' = [1, a']$ en $B = [1, b]$, pongamos

$$c = f(a') \qquad C = B - \{c\}.$$

Designemos por g la aplicación de A en C coincidiendo con f sobre A, es también una inyección; o bien $c = b$ o bien $c \neq b$: si $c = b$, $C = [1, 'b]$ ($'b$ predecesor de b), luego (hipótesis de recurrencia) $a \leq 'b$ y (crecimiento de la aplicación $x \to x'$)

$$a' \leq ('b)' = b.$$

Si $c \neq b$, consideremos la aplicación φ de C sobre $[1, 'b]$ definida por

$$\text{si } c > 1 \begin{cases} 1 \leq x < c & \varphi(x) = x \\ c < x \leq b & \varphi(x) = 'x \end{cases}$$
$$\text{si } c = 1 \quad 1 < x \leq b \quad \varphi(x) = 'x.$$

[illegible] de C sobre $[1, 'b]$, luego φ es inyectiva y $\varphi \circ g$ que [illegible] en $[1, 'b]$ es una inyección (compuesta de dos inyecciones, [illegible] hipótesis de recurrencia) $a \leq 'b$ y como más arriba $a' \leq b$. [illegible] está, pues, demostrado.

[illegible] — *Si a y b son dos enteros naturales no nulos, si existe una biyección* [illegible] *sobre* $[1, b]$, $a = b$.

En efecto, según el teorema 1: $a \leq b$, $b \leq a$, luego $a = b$.

Teorema 1'. — *Si a es un entero natural no nulo, toda inyección de* $A = [1, a]$ *en* A *es una biyección.*

El teorema es verdadero para $a = 1$. Sea $a > 1$, supongamos el teorema verdadero para $'a$ (hipótesis de recurrencia) y sea f una inyección de A en A. Pongamos $c = f(a)$ y $C = A - \{c\}$ y sea g la aplicación de $[1, 'a]$ en C que coincide con f sobre $[1, 'a]$.

Si $c = a$, $C = [1, 'a]$ g es biyectiva (hipótesis de recurrencia), luego f también.

Si $c \neq a$, consideremos, como en la demostración del teorema 1 (haciendo $b = a$), la biyección φ de C sobre $[1, 'a]$, $\varphi \circ g$ es entonces una inyección de $[1, 'a]$ en $[1, 'a]$, luego una biyección (hipótesis de recurrencia); y asimismo g y, por tanto, f: con lo que el teorema queda demostrado.

Teorema 2. — *Toda parte no vacía de* **N** *es finita si y sólo si es acotada superiormente.*

a) Toda parte no vacía finita F de N es acotada superiormente. Existe, pues, un $n \neq 0$ y una biyección f de $[1, n]$ sobre F. Razonemos por recurrencia: si $n = 1$, $F = \{f(1)\}$ y F está acotada superiormente por todos los enteros superiores a $f(1)$. Supongamos cierto el teorema para n y sea M una cota superior de $\{f(1), f(2), ..., f(n)\}$, veamos que también es cierto para n'; en efecto, sup $(M, f(n'))$ es una cota superior de $\{f(1), f(2), ..., f(n')\}$.

b) Toda parte no vacía acotada superiormente F de N es finita.

Por la consecuencia del axioma N_3 (§ 28), F tiene un elemento máximo p. Razonemos por recurrencia sobre p.

Si $p = 0$, $F = \{0\}$ y F es finito, pues existe una biyección (única, por otra parte) de $\{0\}$ sobre $\{1\} = [1, 1]$.

Supongamos el teorema verdadero hasta p. Sea F una parte cuyo máximo es p'; pongamos $X = F - \{p'\}$, el elemento máximo de X es $\leq p$, luego existe una biyección (hipótesis de recurrencia) $f: X \to [1, a]$; consideremos la aplicación g de F sobre $[1, a']$ que coincide con f sobre X tal que $g(p') = a'$, es una biyección de F sobre $[1, a']$, siendo el teorema cierto para p es cierto para p', siendo verdadero para $p = 0$, es cierto cualquiera que sea p máximo elemento de F, luego para toda parte no vacía acotada superiormente de N.

Corolario 1. — *Toda parte* G *de una parte finita* F *de* N *es finita.*

En efecto, si $G \subset F$, toda cota superior de F es una cota superior de G.

Corolario 2. — *Toda intersección* I *de partes finitas de* N *es una parte finita de* N.

Corolario 3. — *La reunión de dos partes finitas de* N *es finita.*

Sean F_1 y F_2 las dos partes finitas de N, sup F_1 y sup F_2 existen, sup (sup F_1, sup F_2) es una cota superior de $F_1 \cup F_2$, luego $F_1 \cup F_2$ parte acotada superiormente, es finita.

Corolario 4. — *Toda parte complementaria respecto a* N *de una parte finita de* N *es infinita. En particular,* N *es infinito.*

En efecto, si F es finito $\complement F$ no es finito, ya que si ambos F y $\complement F$ fuesen finitos serían acotados superiormente, así como $F \cup \complement F = N$, pero N no está acotado superiormente (pues N entonces tendría un elemento máximo, lo que es falso: axioma N_2).

31. Propiedades de los conjuntos finitos

Teorema 3. — *Si a y b son dos enteros naturales distintos de cero, si existe una biyección f de un conjunto* F *sobre $[1, a]$ y una biyección g de* F *sobre $[1, b]$ entonces $a = b$.*

$g \circ f^{-1}$ es una biyección de $[1, a]$ sobre $[1, b]$, luego (corolario del teorema 1) $a = b$.

COROLARIO Y DEFINICIÓN. — *Dados varios conjuntos* F *finitos no vacíos equipotentes entre sí, existe un entero natural único no nulo a tal que cada conjunto* F *sea equipotente a* $[1, a]$.

A a se le llama el cardinal *de* F, *o bien el número de elementos de* F; *se escribe*

$$\text{card } F = a.$$

Por definición card $\varnothing = 0$.

Así card $[1, a] = a$, card $\{x\} = 1$ y si $x \neq y$, card $\{x, y\} = 2$, ...

TEOREMA 4. — *Toda parte* A *de un conjunto finito* B *es finito y* card A $\leq$ card B. *Y* card A = card B *implica* A = B.

Si B = $\varnothing$, lo es también A y el teorema es evidente. Igualmente si A = $\varnothing$. Supongamos A y B no vacíos. Sea g una biyección de B sobre $[1, b]$ [illegible] designemos por g' la aplicación de A sobre g(A) que coincide con g sobre A, es una biyección; pero $g(A)$ parte de $[1, b]$ es finito (corolario [illegible] del teorema [illegible]), A es, pues, finito. Llamemos f una biyección [illegible] (a = card de A) y φ la inyección canónica de A en B (ver § 12, c); [illegible]

$$[1, a] \xrightarrow{f} A \xrightarrow{\varphi} B \xrightarrow{g} [1, b]$$

[illegible] es una biyección, luego (teorema 1) $a \leq b$. [illegible] es una biyección (teorema 1'), también lo es $\varphi = g^{-1} \circ h \circ f$, [illegible] que φ es supràyectiva, luego es la identidad de A y A = B.

[illegible] *La intersección de una familia cualquiera finita o infinita de con-* [illegible] *es finita.*

[illegible] *Si f es una aplicación de* E *finito en un conjunto cualquiera* F, [illegible]

$$\text{card } f(E) \leq \text{card } E.$$

Además, card $f(E)$ = card E *si y sólo si f es inyectiva.*

Perteneciendo y a $f(E)$, escojamos un x único en cada subconjunto $f^{-1}(y)$, [illegible] A la parte de E descrita por estos elementos x; consideremos la aplicación [illegible] A sobre $f(E)$ coincidiendo con f sobre A, g es biyectiva, luego (teorema 4)

$$\text{card } f(E) = \text{card } A \leq \text{card } E.$$

Por otra parte, en virtud del teorema 4, card $f(E)$ = card A = card E equivale a A = E, es decir, a "f es inyectiva".

COROLARIO 3. — *Si f es una aplicación supràyectiva de* E *finito sobre* F, F *es finito y*

$$\text{card } F \leq \text{card } E.$$

Además, card F = card E *si y sólo si f es biyectiva.*

En efecto, es suficiente de aplicar el corolario precedente a F = f(E).

Corolario 4. — *Si f es una aplicación inyectiva de* E *en* F, *si* $f(E)$ *es finito, también lo es* E *y*

$$\text{card } E = \text{card } f(E).$$

En efecto, la aplicación g de E sobre $f(E)$ coincidiendo con f sobre E es una biyección (ver § 16).

Algunos de los resultados precedentes pueden resumirse en el enunciado siguiente:

Corolario 5. — *Dados dos conjuntos* E *y* F *finitos de igual cardinal y siendo f una aplicación de* E *en* F, *las propiedades siguientes son equivalentes:*

a) *f es inyectiva.*
b) *f es suprayectiva.*
c) *f es biyectiva.*

OBSERVACION

Este teorema es falso para los conjuntos infinitos, es suficiente considerar las aplicaciones f y g de **N** en sí mismo definidas por

	$f(n) = 2n$	(inyectiva, no suprayectiva)
n par	$g(n) = n/2$	(suprayectiva, no inyectiva)
n impar	$g(n) = 0.$	

Teorema 5. — *La reunión de dos conjuntos finitos es finita.*

Cualesquiera que sean A y B se tiene

$$A \cup B = A \cup (B - A).$$

Luego se puede suponer A y B disjuntos.

Si A es vacío, $\varnothing \cup B = B$, el teorema es cierto.

Supongamos el teorema demostrado para card $A = a \neq 0$.

Consideremos un conjunto A′ de cardinal a' (siguiente de a), pongamos

$$A' = \{x_1, x_2, \ldots, x_{a'}\} \qquad B = \{y_1, y_2, \ldots, y_b\}$$

existe una biyección f de $(A' - \{x_{a'}\}) \cup B$ sobre $[1, c]$ (hipótesis de recurrencia), sea g una aplicación de $A' \cup B$ en $[1, c']$ definida por

$$[\forall z \in (A' - \{x_{a'}\}) \cup B] \qquad g(z) = f(z) \in [1, c]$$

y

$$g(x_{a'}) = c' \in [1, c].$$

Es ciertamente una biyección, pues f es una biyección y A′ y B son disjuntos. Luego $A' \cup B$ es finito.

Por recurrencia, se demostrará el resultado siguiente:

Corolario 1. — *La reunión de una familia finita de conjuntos finitos es finita.*

Corolario 2. — *Sea f una aplicación de* E *en* F *finito tal que cada imagen recíproca de un elemento y de* $f(E)$ *sea finita, entonces* E *es finito.*

Sea R la relación de equivalencia asociada a f (ver § 19 *b*), la aplicación canónica φ: $E/R \to f(E)$ es biyectiva, luego E/R es finito (pues $f(E)$ parte de F finito es finito). Luego E reunión de un número finito de clases $f^{-1}(y)$ finitas es finito.

Corolario 3. — *El conjunto producto de dos conjuntos finitos es finito.*

Sean A y B dos conjuntos finitos descritos, respectivamente, por x y por y, consideremos la primera proyección pr_1 *de* $A \times B$ sobre A (ver § 12, *d*), si card $A = a$ y card $B = b$, card $(pr_1^{-1}x) =$ card B, pues para x fijo, la aplicación $y \to (x, y)$ es una biyección, luego se puede aplicar el corolario anterior.

32. Notaciones para una familia finita de elementos o de conjuntos

Hemos indicado en el § 17 las notaciones para una familia de elementos o de conjuntos cuyos índices pertenecen a un conjunto I. Si el conjunto I es una parte finita de N, se dice que la familia de elementos, o de conjuntos, es una *familia finita*. Si no es vacía, se puede tomar por conjunto de índices $[1, n]$ y se escribirá

$$(x_i)_{1 \leq i \leq n} \qquad (A_i)_{1 \leq i \leq n}$$

observemos que n no será el cardinal del conjunto de las x_i (o de las A_i) más que en el caso en que la aplicación

$$i \to x_i \qquad (i \to A_i)$$

sea biyectiva.

La intersección, la reunión de los $A_i (1 \leq i \leq n)$ se notarán, respectivamente

$$\bigcup_{i=1}^{i=n} A_i \qquad \bigcap_{i=1}^{i=n} A_i.$$

Análogamente, si para $1 \leq i \leq n$, a_i describe A_i, definiremos el conjunto producto de los conjuntos $(A_i)_{1 \leq i \leq n}$ por recurrencia de 1 a n poniendo

$$A_1 \times A_2 \times \ldots \times A_{p'} = (A_1 \times A_2 \times \ldots \times A_p) \times A_{p'}$$

(p' siguiente de p, siendo p tal que $1 \leq p \leq 'n$) (ver § 29). Escribiremos

$$A_1 \times A_2 \times \ldots \times A_n = \prod_{i=1}^{i=n} A_i.$$

Este conjunto está descrito por $(a_1, a_2, \ldots, a_n)$ que llamaremos una n-étupla. a_i se llama el n-coordenado de la n-étupla.

Observemos que en las notaciones precedentes

$$\bigcup_{i=1}^{i=n} A_i \quad \bigcup_{i=1}^{i=n} A_i \quad \prod_{i=1}^{i=n} A_i.$$

i representa un índice entero que toma todos los valores de 1 *a n. Podemos reemplazarlo por cualquier letra j, k,* ... *o símbolo, excepto por* **1**; *primer valor especificado, o por n: último valor especificado.* Así, por ejemplo,

$$\bigcup_{i=1}^{i=n} A_i = \bigcup_{p=1}^{p=n} A_p$$

contrariamente un símbolo como $\bigcup A_n$ no tiene sentido. Diremos que *i*, o *p*, es un índice *mudo*, o que los términos precedentes *no contienen i* (ni *p*) (ver § 6, *b*, nota 2).

Supongamos, además, que $A_1, \ldots, A_n$ sean conjuntos finitos, hemos visto en el § 31 (corolario 1 del teorema 4) que toda intersección finita o infinita de conjuntos finitos era finita, y que toda reunión finita de conjuntos finitos era finita (corolario 1 del teorema 1), se demuestra por recurrencia mediante el corolario 3 el resultado siguiente:

COROLARIO 4. — *El conjunto producto de un número finito de conjuntos finitos es un conjunto finito.*

III. Operaciones con los enteros naturales

33. Sumas de dos o varios enteros naturales

a) TEOREMA Y DEFINICIÓN. — *Sean dos conjuntos finitos disjuntos* A *de cardinal a y* B *de cardinal b y* A′ *y* B′, *dos conjuntos disjuntos equipotentes, respectivamente, a* A *y a* B, *entonces*

$$\text{card } (A \cup B) = \text{card } (A' \cup B').$$

El cardinal de $A \cup B$ *se llama la* suma *de los enteros a y b y se designa* $a + b$; *a y b son los* términos *de la suma; la operación que permite pasar del par* (a, b) *a* $a + b$ *se llama* suma *en* **N**.

Existe, en efecto, una biyección *f* de A sobre A′ y una biyección g de B sobre B′; consideremos la aplicación

$$F: A \cup B \to A' \cup B'$$

definido por

$$(\forall x \in A) \quad F(x) = f(x); \qquad (A\, x \in B) \quad F(x) = g(x)$$

es una biyección, cualquiera que sea *y* de $A' \cup B'$, la ecuación $F(x) = y$ tiene una única solución, al ser A′ y B′ disjuntos y *f* y g biyecciones.

Por otra parte, A ∪ B es finito (ver § 31, t. 5), puesto que card (A ∪ B) no depende de A y de B, sino del card A y card B, podemos escribir por definición

$$(\text{card A}) + (\text{card B}) = \text{card}(A \cup B).$$

Las propiedades de la reunión (ver § 5) muestran (con las notaciones evidentes) que cualesquiera que sean los enteros naturales, a, b, c cardinales respectivos de conjuntos finitos dos a dos disjuntos

$(a + b) + c = a + (b + c)$ (asociativa)
$a + b = b + a$ (conmutativa)
$a + 0 = 0 + a = a$ (existencia de un elemento neutro para la suma).

Por inducción sobre el entero $i(1 \leq i \leq n)$, definiremos la suma de n enteros representada

$$a_1 + a_2 + \ldots + a_n = \sum_{i=1}^{i=n} a_i.$$

Observemos que en el término del segundo miembro i es *mudo* (§ 6, *b*, nota 2). Las demostraciones fáciles en su principio, pero pesadas y que omitiremos (ver capítulo 3, ejer. 42 y 43), establecen que en una suma de enteros se puede modificar arbitrariamente el orden de los términos de la suma, y reemplazar un grupo arbitrario de estos términos por su suma, sin modificar la suma primitiva.

b) Siendo la suma de dos enteros un entero bien determinado, cualesquiera que sean a, b, c, tenemos

$$a = b \Rightarrow a + c = b + c$$

y de una manera más general

$$(i = 1, 2, \ldots, n) \qquad a_i = b_i \qquad \sum_{i=1}^{i=n} a_i = \sum_{i=1}^{i=n} b_i,$$

es decir: se puede "sumar miembro a miembro" un número finito de igualdades entre enteros.

Recíprocamente, sean tres conjuntos finitos A, B, C (de cardinales respectivos a, b, c), se ve fácilmente que

$$\left\{ \begin{array}{c} A \cap C = B \cap C = \varnothing \\ \text{y} \\ A \cup C = B \cup C \end{array} \right. \Rightarrow A = B$$

luego cualesquiera que sean los enteros naturales a, b, c

$$a + c = b + c \Rightarrow a = b;$$

se dice que todo entero natural es *regular* para la suma, o también que las igualdades entre sumas se pueden *simplificar* por un mismo número entero.

Por otra parte, hemos visto (ver § 31, teorema 4)

$$A \subset B \Rightarrow a \leq b$$

se ve fácilmente que

$$\left\{\begin{array}{l} A \cap C = B \cap C = \varnothing \\ \qquad\quad \text{y} \\ A \subset B \end{array}\right. \Leftrightarrow A \cup C \subset B \cup C$$

luego cualesquiera que sean los enteros naturales a, b, c

$$a \leq b \Leftrightarrow a + c \leq b + c;$$

se dice que la relación de orden $a \leq b$ definida sobre **N** es *compatible* con la suma en **N**; se deduce que se puede "sumar miembro a miembro" un número finito de desigualdades del mismo sentido.

c) Busquemos el siguiente de a. Si $a = 0$, $0' = 1 = 0 + 1$; supongamos $a \neq 0$, $a = \text{card } [1, a]$ y $a' = \text{card } [1, a']$, pero por definición del siguiente, $]a, a'[$ es vacío, luego

$$[1, a'] = [1, a] \cup \{a'\} \qquad \text{y} \qquad [1, a] \cap \{a'\} = \varnothing$$

luego $a' = a + 1$.

Dados dos enteros naturales a y b, si existe un entero natural d tal que $b + d = a$, es único (regularidad de todo entero para la suma).

Podemos considerar b y d como los cardinales respectivos de B y D con $B \cap D = \varnothing$, si ponemos $A = B \cup D$, $a = \text{card } A$, tendremos (ver § 31, teorema 4)

$$A \supset B \qquad a \geq b.$$

Supongamos, pues, $a \geq b$, si $a = b$, $D = \varnothing$ y $d = 0$; por otra parte, si $b = 0$, $d = a$; supongamos, pues, $0 < b < a$, tendremos entonces

$$[1, a] = [1, b] \cup [b', a] \qquad \text{y} \qquad [1, b] \cap [b', a] = \varnothing,$$

de donde

$$a = b + \text{card } [b', a],$$

luego d existe, es el card $[b', a]$, se le llama la *diferencia* de a y b y se designa $a - b$, la operación $(a, b) \rightarrow a - b$ es la *sustracción*, no está definida más que cuando $a \geq b$.

Si $a \neq 0$ tenemos, como más arriba,

$$[1, a] = [1, 'a] \cup \{a\},$$

de donde $'a = a - 1$.

En adelante sólo emplearemos las notaciones $a + 1$ y $a - 1$ en lugar de a' y $'a$.

EJERCICIOS

1. Demostrar que sobre el conjunto de pares de enteros, en que está definida la sustracción, la relación entre (a_1, b_1) y (a_2, b_2) dada por

$$a_1 - b_1 = a_2 - b_2$$

es una relación de equivalencia.

2. Demostrar que la relación, definida sobre $\mathbf{N} \times \mathbf{N}$, entre (a_1, b_1) y (a_2, b_2)

$$a_1 + b_2 = a_2 + b_1$$

es una relación de equivalencia.

14. Multiplicación de dos o más enteros naturales

a) TEOREMA Y DEFINICIÓN. — *Sean dos conjuntos finitos* A *de cardinal* a *y* B *de cardinal* b *y* A′ *y* B′ *dos conjuntos, respectivamente, equipotentes a* A *y a* B, *entonces*

$$\text{card } (A \times B) = \text{card } (A' \times B').$$

El cardinal de $A \times B$ *se llama el* producto *de los enteros naturales* a *y* b *y se escribe* $a \cdot b$ *o* ab*;* a *y* b *son los* factores *del producto; la operación que permite pasar del par* (a, b) *a* ab *se llama* multiplicación *en* $\mathbf{N}$.

Existe, en efecto, una biyección f de A sobre A′ y una biyección g de B sobre B′; consideremos la aplicación

$$h: \ A \times B \rightarrow A' \times B'$$

definida por

$$(\forall x \in A) \quad (\forall y \in B) \qquad h(x, y) = (f(x), g(y)).$$

Que es evidentemente una biyección, pues cualquiera que sea el par (x', y') de $A' \times B'$

$$(f(x), g(y)) = (x', y') \Leftrightarrow [f(x) = x' \quad \text{y} \quad g(y) = y'],$$

pero estas dos últimas ecuaciones tienen una solución única cualquiera que sea x' de A′ e y' de B′, luego h es una biyección.

Por otra parte, $A \times B$ es finito (ver § 31, corolario 3 del teorema 5), card $(A \times B)$ no dependiendo de A y de B, sino únicamente del card A y card B; podemos poner por definición

$$(\text{card } A)(\text{card } B) = \text{card } (A \times B).$$

La definición de $A \times B \times C$ (ver § 8) muestra que, cualesquiera que sean los enteros naturales $a, b, c,$

$$(ab)c = a(bc) \qquad \text{(asociatividad)}.$$

La aplicación de $A \times B$ en $B \times A$ definida por

$$(x, y) \rightarrow (y, x)$$

(con las notaciones del principio del párrafo) es una biyección, luego cualesquiera que sean los enteros naturales a y b

$$ab = ba \qquad \text{(conmutatividad).}$$

Por otra parte, x describiendo A y siendo y fija, la aplicación de A en $A \times \{y\}$ definida por

$$x \rightarrow (x, y)$$

es una biyección, luego para todo entero natural a

$$a = a1 = 1a \qquad \text{(existencia de un elemento neutro para la multiplicación).}$$

En fin, si B_1 y B_2 son dos conjuntos finitos y disjuntos, análogamente lo son $A \times B_1$ y $A \times B_2$ y (ver § 8)

$$A \times (B_1 \cup B_2) = (A \times B_1) \cup (A \times B_2)$$

luego, cualesquiera que sean los enteros naturales a, b_1, b_2,

$$a(b_1 + b_2) = ab_1 + ab_2 \qquad \text{(distributividad de la multiplicación con relación a la suma).}$$

Por recurrencia sobre el entero $i(1 \leq i \leq n)$ definiremos el producto de n enteros que se escribe

$$a_1, a_2, \ldots, a_n = \prod_{i=1}^{i=n} a_i.$$

Observemos que, en el término del segundo miembro, i es *mudo* (ver § 6, *b*, nota 2).

Como para la suma, admitiremos que en un producto de enteros se puede modificar arbitrariamente el orden de los factores y reemplazar un grupo arbitrario de factores por su producto, sin modificar el producto primitivo (ver (capítulo 3, ej. 42 y 43).

Por recurrencia sobre $j(1 \leq j \leq q)$, se mostrará que

$$a\left(\sum_{j=1}^{j=q} b_j\right) = \sum_{j=1}^{j=q} ab_j$$

y seguidamente por recurrencia sobre $i(1 \leq i \leq p)$

$$\left(\sum_{i=1}^{i=p} a_i\right)\left(\sum_{j=1}^{j=q} b_j\right) = \sum_{i=1}^{i=p}\sum_{j=1}^{j=q} a_i b_j.$$

b) Las propiedades del producto cartesiano indicadas en el § 8 muestran que, para todo entero natural a,

$$a0 = 0a = 0$$

y

$$ab = 0 \Rightarrow (a = 0 \quad \text{o} \quad b = 0).$$

Por otro lado, siendo el producto de dos enteros un entero bien determinado, cualesquiera que sean a, b, c,

$$a = b \Rightarrow ac = bc$$

y más generalmente

$$(i = 1, 2, \ldots, n) \quad a_i = b_i \Rightarrow \prod_{i=1}^{i=n} a_i = \prod_{i=1}^{i=n} b_i$$

recíprocamente (ver § 8)

$$(\mathrm{C} \neq \varnothing \quad \text{y} \quad \mathrm{A} \times \mathrm{C} = \mathrm{B} \times \mathrm{C}) \Rightarrow \mathrm{A} = \mathrm{B}$$

luego, cualesquiera que sean a y b y cualquiera que sea $c \neq 0$,

$$ac = bc \Rightarrow a = b.$$

Se dice que *todo entero natural no nulo es regular para la multiplicación,* o bien que se puede *simplificar* por un mismo entero natural no nulo las igualdades entre productos.

En fin (ver § 8),

$$\mathrm{A} \subset \mathrm{B} \Rightarrow \mathrm{A} \times \mathrm{C} \subset \mathrm{B} \times \mathrm{C}$$

luego

$$a \leq b \Rightarrow ac \leq bc,$$

en cambio se tendrá solamente

$$(ac \leq bc \quad \text{y} \quad c \neq 0) \Rightarrow a \leq b.$$

Se dice que la relación $a \leq b$ es *compatible con la multiplicación por un entero natural no nulo.*

c) OBSERVACION

Consideremos la suma $\sum_{i=1}^{i=a} b_i$ de a enteros todos iguales a b, se ve fácilmente por inducción a partir de $a = 1$ y utilizando la distributividad de la multiplicación respecto a la suma que

$$\sum_{i=1}^{i=a} b_i = ab.$$

Es ésta la definición que se toma algunas veces para el producto de dos enteros naturales: la misma hace desempeñar papeles diferentes a a y a b («multiplicando», multiplicador»), de donde surgen las dificultades para la demostración de la conmutatividad: por esto hemos preferido la definición simétrica que hemos dado.

d) Dados dos enteros naturales a y b ($b \neq 0$), si existe q tal que $a = bq$ es único (regularidad de todo entero no nulo para la multiplicación). Se dice que a es un *múltiplo* de b, o que b *divide* a; esta última relación se escribe $b \mid a$; hemos visto en el § 22 que era una *relación de orden parcial* definida sobre $\mathbf{N}^*$.

Cuando existe q, se le llama el cociente de a por b y se escribe a/b. Se escribirá $b\mathbf{N}$ el conjunto de los enteros naturales *múltiples* de b; así $0\mathbf{N} = \{0\}$, $1\mathbf{N} = \mathbf{N}$, $2\mathbf{N}$ es el conjunto de los enteros naturales pares (ver § 50).

EJERCICIOS

1. Demostrar que si $b_1 - b_2$ existe, también existe $ab_1 - ab_2$ y que

$$a(b_1 - b_2) = ab_1 - ab_2.$$

2. Demostrar que la relación $a_1b_2 = a_2b_1$ entre los pares (a_1, b_1) y (a_2, b_2) de $\mathbf{N} \times \mathbf{N}^*$ es una relación de equivalencia.

3. La relación $b \mid a$, ¿es compatible con la adición, con la multiplicación en $\mathbf{N}^*$?

35. Aplicaciones de la suma y de la multiplicación de enteros a los cardinales de conjuntos finitos

Sean dos conjuntos finitos A y B no forzosamente disjuntos; pongamos

$$A' = A - (A \cap B) \qquad B' = B - (A \cap B)$$

tendremos (al ser A' y $A \cap B$ finitos y disjuntos, igualmente que B' y $A \cap B$)

$$\text{card } A' = \text{card } A - \text{card } (A \cap B) \qquad \text{card } B' = \text{card } B - \text{card } (A \cap B).$$

Por otro lado,

$$A \cup B = A' \cup B' \cup (A \cap B)$$

los tres conjuntos A', B', $A \cap B$ son disjuntos dos a dos, de donde

$$\text{card } (A \cup B) = \text{card } A + \text{card } B + \text{card } (A \cap B),$$

es decir, $\text{card } (A \cup B) \leq \text{card } A + \text{card } B$, verificándose la igualdad si y sólo si $\text{card } (A \cap B) = 0$, es decir, si A y B son disjuntos.

Por recurrencia podemos extender este resultado a n conjuntos finitos $F_1, F_2, \ldots, F_n$ $(n \geq 2)$, considerando los dos conjuntos $(F_1 \cup F_2 \ldots \cup F_{n-1})$ y F_n, de donde

TEOREMA. — *Dados n conjuntos finitos $F_i (1 \leq i \leq n)$, se tiene*

$$\text{card}\left(\bigcup_{i=1}^{n} F_i\right) \leq \sum_{i=1}^{n} \text{card } F_i.$$

Se verifica la igualdad si y sólo si los F_i son disjuntos dos a dos.

COROLARIO 1. — *Si (F_i) $1 \leq i \leq n$ es un recubrimiento finito de una parte finita A de un conjunto E, siendo las partes F_i de E finitas, se tiene*

$$\text{card } A \leq \text{card} \bigcup_{i=1}^{n} F_i \leq \sum_{i=1}^{n} \text{card } F_i.$$

COROLARIO 2. — *Si* $(F_i)_{(1\leq i\leq n)}$ *es una partición finita de un conjunto* E, *si las partes* F_i *de* E *son finitas, se tiene*

$$\text{card } E = \sum_{i=1}^{n} \text{card } F_i.$$

COROLARIO 3. — *Si* f *es una aplicación de* E *en* F *finita tal que cada subconjunto de* E, $f^{-1}(y)$ *sea finito, siendo* n *el cardinal de* $f(E)$, *se tiene que*

$$\text{card } E = \sum_{i=1}^{n} \text{card } f^{-1}(y_i).$$

En efecto, $f(E)$ es finito; sea n su cardinal, $(f^{-1}(y_i))\,(1 \leq i \leq n)$ es una partición de E.

Aplicando la observación del § 34, *c*, obtendremos el resultado siguiente:

COROLARIO 4. *(Principio de los pastores.)* — *Si* f *es una suprayección de* E *sobre* F *finito de cardinal* n *tal que, cualquiera que sea* y *de* F, card $f^{-1}(y) = p$, tenemos

$$\text{card } E = np.$$

EJERCICIO

Dados p elementos repartidos entre n conjuntos de una familia, si $p > n$ existe al menos un conjunto de la familia conteniendo al menos dos elementos (*principio de* [illegible]).

35. División euclídea de los enteros naturales

a) Sean a y b dos enteros naturales ($b \neq 0$), consideremos la aplicación f de **N** en b**N** definida por $x \to f(x) = bx$; es suprayectiva por definición, es inyectiva, pues b no nulo es regular para la multiplicación; es, pues, una biyección de **N** sobre b**N**; b**N** no está acotado superiormente, si no sería finito (ver § 30, teorema 2), cualquiera que sea a, existe, pues, p tal que $bp > a$, de donde:

TEOREMA Y DEFINICIÓN. — *Cualesquiera que sean el entero natural* a *y el entero natural* $b \neq 0$, *existe un entero natural* p *tal que* $bp > a$; *se dice que el orden total definido sobre* **N** *es* arquimediano *con respecto a la suma.*

Se puede mostrar un tal entero; por ejemplo, $a + 1$; en efecto,

$$(a + 1)b = ab + b \geq a + b \geq a + 1 > a.$$

b) Sea P el conjunto de los enteros p tales que $pb > a$; acabamos de ver que este conjunto no es vacío; tiene, pues, un mínimo (ver § 28, axioma N_1); este elemento n no es nulo (pues $0b = 0 \leq a$); se puede, pues, escribir

$q+1$; qb es, pues, inferior o igual a a (puesto que $q+1$ es el más pequeño entero tal que $pb > a$), luego existe q único tal que

$$bq \leq a < b(q+1)$$

pongamos $r = a - bq$; r es único y $0 \leq r < b$, luego:

TEOREMA Y DEFINICIÓN. — *A todo par de enteros naturales a y b tal que $b \neq 0$, se puede hacer corresponder un par único de enteros naturales q y r tales que*

$$a = bq + r \quad \text{y} \quad 0 \leq r < b.$$

La operación que permite pasar del par (a, b) al par (q, r) se llama división euclídea *de los enteros naturales, q y r son, respectivamente, el* cociente *y el* resto *en esta división.*

Si $r = 0$, b divide a y q es entonces el *cociente* en **N** de a por b.
Observemos que si $a < b$, $q = 0$ y $r = a$.

37. Exponencial de base a y de exponente b

Cualquiera que sea el entero natural a y cualquiera que sea el entero natural $b \neq 0$, escribiremos

(1) $$a^1 = a, \quad a^2 = a^1 a, \quad \ldots, \quad a^b = a^{b-1}a,$$

en particular para todo $b \neq 0$: $0^b = 0$.
Escribiremos también para todo $a \neq 0$

(2) $$a^0 = 1.$$

Estas igualdades muestran que para todo par $(a, b) \neq (0, 0)$ el símbolo a^b está definido, de donde:

DEFINICIÓN. — *La aplicación de $\mathbf{N} \times \mathbf{N} - \{(0, 0)\}$ en $\mathbf{N}$ definida por*

$$(a, b) \rightarrow a^b$$

se llama exponencial *de* base *a y de* exponente *b.*

Se verificará sin esfuerzo cuando todos los símbolos utilizados tienen un sentido

$$a^p a^q = a^{p+q}, \qquad a^p b^p = (ab)^p, \qquad (a^p)^q = a^{pq},$$

pero esta operación no es ni *conmutativa* ni *asociativa*; se tiene, por ejemplo,

$$2^3 = 8 \qquad\qquad 3^2 = 9$$
$$(2^3)^2 = 8^2 = 64 \qquad 2^{(3^2)} = 2^9 = 512$$

por convención a^{b^c} representa siempre $a^{(b^c)}$ (en efecto, la notación $(a^b)^c$ es inútil, puesto que $(a^b)^c = a^{bc}$).

EJERCICIOS

1. Demostrar que en **N** es

$$a \leq b \Leftrightarrow a^n \leq b^n, \qquad a < b \Leftrightarrow a^n < b^n, \qquad (\text{si} \quad n \neq 0).$$

2. Demostrar que en **N**, para todo $a > 1$ y todo $n \geq 1$: $a^n > 1 + na$.

3. Deducir del ejercicio precedente que en **N**, para todo $a > 1$ y para todo b existe n tal que: $a^n > b$ (se dice que el orden total definido sobre **N** es arquimediano para la multiplicación).

11. Conclusión

Cualquiera que sea la pareja (a, b) de enteros naturales (excepto la pareja $(0, 0)$ para la exponencial), hemos definido tres operaciones que tienen por resultados respectivos

$$a + b, \qquad ab, \qquad a^b,$$

pero las ecuaciones

$$\begin{array}{llll} (1) & b + x = a & (2) \quad b \neq 0 & bx = a \\ (3) & b^x = a & (3') & x^b = a \end{array}$$

no tienen siempre soluciones en **N**, para (1) y (2) la condición de existencia y de unicidad de x es, respectivamente, $b \leq a$ y $b \mid a$.

Uno de los fines de algunos estudios de los capítulos siguientes será el construir conjuntos conteniendo **N** y en los que estas ecuaciones tendrán siempre soluciones: para (1) éste será **Z**, para (2) lo será **Q**, para (3) y (3′) lo será **R**.

Además, (1) tendrá siempre soluciones en **Z**, incluso si a y b no se toman en **N**, sino en **Z**; análogamente (2) tendrá siempre soluciones en **Q**, al tomar a y b en **Q** $(b \neq 0)$.

Por el contrario, si queremos dar un sentido a la ecuación (3) para a y b reales, será necesario suponer a y b estrictamente positivos. En fin, para resolver ciertas ecuaciones (3′) tales como $x^2 = -1$, nos veremos obligados a introducir un nuevo conjunto **C**, conjunto de los números complejos.

Cada conjunto introducido en el orden indicado será un superconjunto del precedente

$$\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R} \subset \mathbf{C}.$$

Sobre cada uno de estos conjuntos se definirán una suma, una multiplicación, que *prolongan* las operaciones ya definidas sobre el o los conjuntos anteriores.

IV. Análisis combinatorio

39. Número de aplicaciones de un conjunto finito en un conjunto finito

Sean dos conjuntos M y N finitos no vacíos de cardinales respectivos m y n.

Designamos por $\mathfrak{F}(M, N)$ el conjunto de las aplicaciones de M en N.

Si $m = 1$, $M = \{a\}$ y $N = \{b_1, b_2, \ldots, b_n\}$, siendo biyectiva la aplicación $i \to b_i$, las aplicaciones de M en N están definidas por

$$a \to f_i(a) = b_i \qquad (1 \leq i \leq n),$$

luego $\mathfrak{F}(M, N)$ es finito y tiene por cardinal n.

Supongamos que para $m = p - 1$, $\mathfrak{F}(M, N)$ sea finito (hipótesis de recurrencia); sea $m = p$ y a un elemento de M; pongamos

$$M' = M - \{a\} \qquad M'' = \{a\}.$$

A toda aplicación f de M en N hacemos corresponder su *restricción* $f_{M'}$ a M' mediante la aplicación

$$F: \ \mathfrak{F}(M, N) \to \mathfrak{F}(M', N),$$

es decir,

$$f_{M'} = F(f).$$

$F^{-1}(f_{M'})$ es el conjunto de las *prolongaciones* de $f_{M'}$, cada uno está determinado por $f(a) \in N$ y son disjuntos dos a dos; luego

$$\text{card } F^{-1}(f_{M'}) = \text{card } N = n.$$

En consecuencia, $\mathfrak{F}(M, N)$ está definido (ver § 35, corolario 4) y si ponemos $\varphi(p, n) = \text{card } \mathfrak{F}(M, N)$; cuando card $M = p$, tenemos

$$\begin{aligned}
&\varphi(1, n) = n \\
&\varphi(2, n) = n \quad \varphi(1, n) \\
&\cdot \\
&\cdot \\
&\cdot \\
&\varphi(p, n) = n \quad \varphi(p - 1, n) \\
&\cdot \\
&\cdot \\
&\cdot \\
&\varphi(m, n) = n \quad \varphi(m - 1, n)
\end{aligned}$$

multiplicando estas igualdades miembro a miembro, no siendo ningún factor nulo, tendremos

$$\varphi(m, n) = n^m.$$

TEOREMA. *Siendo* M *y* N *no vacíos y de cardinales respectivos* m *y* n, *el conjunto de aplicaciones* $\mathfrak{F}(M, N)$ *es finito y tiene por cardinal* n^m.

OBSERVACIÓN

Es a causa de esta fórmula que se representa algunas veces $\mathfrak{F}(M, N)$ por N^M, cualquiera que sean los conjuntos M y N.

[illegible] Número de las inyecciones de un conjunto finito en un conjunto finito. Variaciones. Permutaciones

a) Notemos por $\mathfrak{J}(M, N)$ el conjunto de las inyecciones de M finito no vacío en N finito no vacío; $\mathfrak{J}(M, N)$ es finito, puesto que es una parte de $\mathfrak{F}(M, N)$ ([illegible] 30).

Siendo f una inyección de M en N, se tiene

$$\text{card } M = \text{card } f(M) \leq \text{card } N,$$

puesto que $f(M)$ está contenido en N, luego $m \leq n$.

[illegible] las mismas notaciones que en el párrafo precedente poniendo [illegible] (M, N) cuando card $M = p$.

[illegible] de f de M en N hacemos corresponder su *restricción* [illegible] por la aplicación

$$\theta: \mathfrak{J}(M, N) \to \mathfrak{J}(M', N).$$

[illegible] $\theta^{-1}(i_{M'})$ es el conjunto de las *prolongaciones* de $i_{M'}$, cada [illegible] por $i(a)$, pero $i(a)$ no se debe tomar en N, sino en [illegible] i siendo inyectiva, $i(a)$ no puede tener un valor ya tomado [illegible]

$$\text{card } \theta^{-1}(i_{M'}) = \text{card } N - \text{card } i_{M'}(M') = n - (p - 1),$$

siendo $i_{M'}$ inyectiva: card $i_{M'}(M') = $ card (M') (§ 31, corolario 4 de [illegible] 4), luego (principio de los pastores, § 35)

$$\text{card } \mathfrak{J}(M, N) = (n - p + 1) \text{ card } \mathfrak{J}(M', N),$$

[illegible] otro lado, $\psi(1, n) = n$, de donde

$$\begin{aligned}
\psi(1, n) &= n \\
\psi(2, n) &= (n - 1) \quad \psi(1, n) \\
&\vdots \\
\psi(p, n) &= (n - p + 1) \quad \psi(p - 1, n) \\
&\vdots \\
\psi(m, n) &= (n - m + 1) \quad \psi(m - 1, n)
\end{aligned}$$

multiplicando estas igualdades miembro a miembro, no siendo nulo ninguno de los factores $\psi(p, n)$, pues $m \leq n$ implica $n - p + 1 > 0$ para $1 \leq p \leq m$, obtendremos

$$\psi(m, n) = n(n-1)\ldots(n-m+1),$$

es decir, el producto de m enteros consecutivos estrictamente decrecientes a partir de n. Obtendremos una expresión más simple gracias a la definición siguiente:

DEFINICIÓN. — *Si n es un entero natural, se designa $n!$ (que se enuncia «n factorial») el entero natural definido por*

$$0! = 1 \quad y \text{ para } \quad n \geq 1 \quad n! = (n-1)!\,n.$$

Luego para $n \geq 1$

$$n! = 1.2.3 \ldots n.$$

La aplicación $n \to n!$ es *estrictamente creciente*[9] para $n \geq 2$; se tiene

$$1! = 1 \quad 2! = 2 \quad 3! = 6 \quad 4! = 24 \quad 5! = 120 \ldots \quad 10! = 3.628.880$$

mediante esta notación es

$$\psi(m, n) = n(n-1)\ldots(n-m+1) = \frac{n!}{(n-m)!}.$$

TEOREMA. — *Si M y N son dos conjuntos finitos no vacíos de cardinales respectivos m y n ($m \leq n$), el conjunto de las inyecciones de M en N es finito y tiene por cardinal*

$$n(n-1)\ldots(n-m+1) = \frac{n!}{(n-m)!}.$$

b) Si $M = \{1, 2, \ldots, m\}$ y $N = \{1, 2, \ldots, n\}$, sea i una inyección de M en N

$$p \to i(p) = i_p$$

la imagen de i, $i(M) = \{i_1, i_2, \ldots, i_m\}$, donde los elementos $i_1, i_2, \ldots, i_m$ están colocados en este orden se llama una *variación sin repetición de los n enteros* 1, 2, ..., n, *m a m*.

De una manera más general si $N = \{x_1, x_2, \ldots, x_n\}$, $i(M) = \{x_{i_1}, x_{i_2}, \ldots, x_{i_m}\}$ con los elementos $x_{i_1}, x_{i_2}, \ldots, x_{i_m}$ colocados en este orden, se llama una *variación sin repetición de los n objetos $x_1, x_2, \ldots, x_n$, m a m.*

Se designa algunas veces el *número de estas variaciones* por A_n^m; es el número de inyecciones de M en N, luego

$$A_n^m = n(n-1)\ldots(n-m+1) = \frac{n!}{(n-m)!}.$$

c) *Caso particular*: card M = card N = n.

(9) 20! tiene 19 cifras en el sistema decimal; se demuestra en Análisis la fórmula asintótica siguiente (llamada fórmula de STIRLING)

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + 0\left(\frac{1}{n}\right)\right).$$

El corolario 5 del teorema 4 (§ 31) indica que toda inyección de M en N es una biyección; según el resultado precedente, hay $n!$, puesto que $(n - n)! = 0! = 1$.

En particular:

TEOREMA. *Hay $n!$ biyecciones de un conjunto* E *con n elementos sobre sí mismo.*

Estas biyecciones se los llama también *permutaciones* de E (ver más adelante § 44). Por abuso de lenguaje si $E = \{x_1, x_2, \ldots, x_n\}$ se llama algunas veces *permutación* la imagen $f(E)$, es decir,

$$f(E) = \{x_{i_1}, x_{i_2}, \ldots, x_{i_n}\}$$

con los elementos $x_{i_1}, \ldots, x_{i_n}$ colocados en este orden.

EJERCICIOS

1. [illegible] cuatro conjuntos A, B, A′, B′, tales que card A = card A′ = a, card B = card [illegible] el número de biyecciones de A × B sobre A′ × B′.

[illegible] de partes (A, B), (A′, B′) de un conjunto finito E de cardinal n [illegible] = A′ ∪ B′, A ∩ B = A′ ∩ B′ = ∅, card A = card A′ = a, [illegible] encontrar el número de biyecciones f de E tal que [illegible] el ejercicio [illegible] al final del capítulo I, y el ejercicio pre-

[illegible]

[illegible] partes de m elementos de un conjunto E con n elementos. [illegible] sin repeticiones

[illegible] un conjunto finito tal que card E = $n \geq 1$, busquemos el número [illegible] con m elementos, naturalmente $m \leq n$. Anotemos $\mathfrak{F}_m(E)$ el con[illegible] estas partes de m elementos.

[illegible] una biyección f de $N = [1, n]$ sobre E, la aplicación de $\mathfrak{F}(N)$ en [illegible] por (siendo A una parte de N) $A \to f(A)$ es biyectiva, luego $\mathfrak{F}(N)$ [illegible] son equipotentes; supondremos, pues, $E = N = [1, n]$.

[illegible] toda inyección i de $M = [1, m]$ en $N = [1, n]$ corresponde una parte [illegible] de N

$$A = i(M) = \{i_1, i_2, \ldots, i_m\}$$

Definimos así una aplicación f de $\mathfrak{J}(M, N)$ en $\mathfrak{F}_m(N)$

$$i \to f(i) = A$$

[illegible] es suprayectiva, pues si $B = \{j_1, j_2, \ldots, j_m\}$ es una parte de N con m elementos, [illegible] aplicación $h \to j_h$ es una inyección y $B = f(j)$. Determinemos $f^{-1}(A)$, es decir, [illegible] las inyecciones de M en N que dan una parte A de N; tendremos

$$i(M) = \{i_1, i_2, \ldots, i_m\} = i'(M) = \{i'_1, i'_2, \ldots, i'_m\}$$

si y sólo si la aplicación de A en sí mismo definida por

$$i_h \to i'_h$$

es biyectiva, pero hay $m!$ biyecciones de A sobre sí mismo (§ 40, c), luego

$$\text{card } f^{-1}(A) = m!$$

y, en consecuencia (principio de los pastores, § 35),

$$\text{card } \mathfrak{J}(M, N) = (m! \text{ card } \mathfrak{F}_m(N)),$$

de donde

$$\boxed{\text{card } \mathfrak{F}_m(N) = \frac{n(n-1)\dots(n-m+1)}{m!} = \frac{n!}{m!(n-m)!}}$$

El razonamiento precedente supone $m > 0$; ahora bien, para $m = 0$, existe una sola parte ($\emptyset$) y la fórmula es también válida.

Este número se representa por la notación $\binom{n}{m}$ o C_n^m, se llama *coeficiente binomial de índices n y m*, veremos más adelante por qué (§ 92).

TEOREMA. — *El número de partes con m elementos de un conjunto de n elementos es igual a*

$$\binom{n}{m} = C_n^m = \frac{n(n-1)\dots(n-m+1)}{m!} = \frac{n!}{m!(n-m)!}.$$

(*Observar* los lugares diferentes de m y n en las dos notaciones $\binom{n}{m}$ y C_n^m)

Toda parte con m elementos de un conjunto con n elementos se llama *combinación sin repetición de n elementos tomados m a n.*

b) La fórmula precedente muestra que

$$C_n^0 = C_n^n \qquad C_n^m = C_n^{n-m}$$

Por otra parte, un cálculo fácil da para $m \geq 1$

$$C_n^m = C_{n-1}^{m-1} + C_{n-1}^m,$$

de ahí la construcción por recurrencia de la tabla de coeficientes binomiales, llamada *triángulo de Pascal*,

$n = 0$	1								
$n = 1$	1	1							
$n = 2$	1	2	1						
$n = 3$	1	3	3	1					
⋮									
$n - 1$	1	C_{n-1}^{1}	C_{n-1}^{2}	…	C_{n-1}^{m-1}	C_{n-1}^{m}	…	1	
n	1	C_{n}^{1}	C_{n}^{2}	…		C_{n}^{m}	…	C_{n}^{n-1}	1.

Ejercicios

1. Demostrar la fórmula $C_n^m = C_n^{n-m}$ teniendo en cuenta que la aplicación $A \to \complement_E A$ [illegible] es biyectiva.

2. [illegible] $C_n^m = C_{n-1}^m + C_{n-1}^{m-1}$ aislando un elemento a en E. Demostrar igual- [illegible]

$$C_n^m = C_{n-2}^m + 2C_{n-2}^{m-1} + C_{n-2}^{m-2},$$

[illegible] bien aislando dos elementos a y b de E.

[illegible] que

$$C_n^m = \frac{n}{m} C_{n-1}^{m-1}, \qquad C_n^m = \frac{n-m+1}{m} C_n^{m-1}.$$

[illegible] n fija, estudiar el crecimiento de la aplicación $m \to C_n^m$.
[illegible] n un número primo, demostrar que para $m \neq 0$ y $m \neq n$, C_n^m es divisible [illegible]

[illegible] Demostrar que

$$\sum_{m=0}^{m=n} C_n^m = 2^n.$$

[illegible] la función característica de una parte A de E, se demostrará que la aplicación [illegible] en el conjunto de las funciones características definidas sobre E (ver § 12, *c*) [illegible] es una biyección y se calculará card $\mathfrak{F}(E, \{0, 1\})$.)

Deducir de lo precedente que card $E = n \Rightarrow$ card $\mathfrak{P}(E) = 2^n$ (este resultado ha sido [illegible] en el ejercicio del § 4, ej. 4; propondremos una tercera demostración en el [illegible], ej. 4).

V. Nociones sobre los conjuntos numerables

42. Definición. — *Un conjunto* E *es* estrictamente numerable *si existe una biyección de* E *sobre el conjunto* N *de los enteros naturales. Un conjunto es* numerable *si es finito o estrictamente numerable.*

Teorema 1. — *Para que un conjunto* E *sea estrictamente numerable, es necesario y suficiente que exista una biyección de* E *sobre un conjunto estrictamente numerable* F.

Si φ es la biyección de F sobre N que existe por hipótesis, existe una biyección f de E sobre F, $\varphi \circ f$ es una biyección de E sobre N. Inversamente, si existe una biyección g de E sobre N, $\varphi^{-1} \circ g$ es una biyección de E sobre F.

Corolario 1. — *El conjunto* N′ *de los enteros pares y el conjunto* N″ *de los enteros impares son estrictamente numerables.*

Se define, en efecto, una biyección de N′ sobre N haciendo corresponder a todo entero par $2n$ su mitad, y una biyección de N″ sobre N haciendo corresponder a todo entero impar $2n + 1$ el número entero n.

Corolario 2. — *La reunión de dos conjuntos* E, F *estrictamente numerables, disjuntos es estrictamente numerable.*

En virtud del teorema 1, existe una biyección φ de E sobre N′ y una biyección ψ de F sobre N″. La aplicación f de $E \cup F$ en N que admite φ por restricción a E y ψ por restricción a F es una biyección de $E \cup F$ sobre N.

EJERCICIOS

1. Toda parte de un conjunto numerable es numerable.
2. La reunión de dos conjuntos numerables disjuntos es numerable.
3. La reunión de dos conjuntos numerables es numerable.

Teorema 2. — *El producto cartesiano* $E \times F$ *de dos conjuntos* E, F *estrictamente numerables es estrictamente numerable.*

Si φ es una biyección de E sobre N y ψ una biyección de F sobre N, la aplicación de $E \times F$ en $N \times N$ que a la pareja (x, y) hace corresponder $(\varphi(x), \psi(y))$ es una biyección. Es suficiente, pues, demostrar que $N \times N$ es numerable. Se puede establecer una biyección de $N \times N$ sobre N de esta manera: se colocan las parejas (m, n) de enteros en una tabla infinita de doble entrada, se enumeran las parejas, dando el número 1 a la pareja (0, 0), el número 2 a la pareja (0, 1), el número 3 a la pareja (1, 0) y más generalmente tomando las parejas según las diagonales sucesivas (las parejas de una misma diagonal están caracterizadas por la suma de los elementos de la pareja) y recorriendo cada diagonal de derecha a izquierda (ver ej. 41, fin de este capítulo).

Corolario. — *La reunión de una familia* $\mathfrak{F}$ *estrictamente numerable de conjuntos estrictamente numerables, dos a dos disjuntos, es estrictamente numerable.*

Sea φ la biyección de $\mathfrak{F}$ sobre **N**. Si E es un conjunto de la familia, sea ψ_E la biyección de E sobre **N**.

Si x pertenece al conjunto F, reunión de la familia $\mathfrak{F}$, x pertenece a un conjunto E, y a uno sólo, puesto que son dos a dos disjuntos. Haciendo corresponder a x la pareja $(\varphi(E), \psi_E(x))$, establecemos una biyección de F sobre $\mathbf{N} \times \mathbf{N}$.

EJERCICIOS

4. Demostrar que la aplicación precedente es una biyección.

5. Sea $\mathfrak{F}$ una familia estrictamente numerable de conjuntos, y sea E_n el conjunto que corresponde a n en una biyección de **N** sobre $\mathfrak{F}$.

Se pone

$$E'_n = E_n - \bigcup_{i=1}^{n-1} E_i$$

[illegible]

[illegible] que los conjuntos E'_n son dos a dos disjuntos y que se tiene

$$\bigcup_{n \in \mathbf{N}} E'_n = \bigcup_{n \in \mathbf{N}} E_n$$

[illegible] de una familia estrictamente numerable de conjuntos estricta- [illegible] estrictamente numerable.

[illegible] reunión de una familia numerable de conjuntos numerables es numerable.

Ejercicios

26. Siendo E un conjunto no vacío, demostrar que el conjunto de los axiomas N_1, N_2, N_3 del § 28 es equivalente al conjunto de los axiomas:

N_1' E está totalmente ordenado y tiene un elemento mínimo m.

N_2 Todo elemento x de E tiene un siguiente x'.

N_3' Toda parte X de E tal que

$$m \in X \quad \text{y} \quad (x \in X \Rightarrow x' \in X)$$

es idéntica a E.

27*. Si E es un conjunto no vacío, demostrar que el conjunto de los axiomas N_1, N_2, N_3 del § es equivalente al conjunto de los axiomas:

N_1 E es ordenado y toda parte no vacía de E tiene un elemento mínimo.

N_2' E no tiene elemento máximo.

N_3'' Toda parte acotada superiormente no vacía de E tiene un elemento máximo.

28. Dos conjuntos E y E_1 que verifican los axiomas N_1, N_2, N_3 poseen las propiedades resumidas en el § 28 (según la consecuencia del axioma N_3). Designando por m y m_1 sus primeros elementos respectivos, demostrar para E y E_1 un teorema análogo al teorema del § 29.

Se propone demostrar que E y E_1 son isomorfos para el orden definido en cada uno de ellos que se notará $\leq$.

a) Si existe una aplicación f_a de $[m, a]$ en E_1 tal que $f_a(m) = m_1$ y $f_a(x') = [f_a(x)]'$ para todo x de $[m', a]$ se dirá que a es *regular*. Demostrar que m' es regular y que si a es regular, lo es también a'.

b) Si $a \leq b$, demostrar que la restricción de f_b a $[m, a]$ coincide con f_a sobre $[m, a]$.

c) Considerar, en fin, la aplicación f de E en E_1 que coincide con f_a sobre $[m, a]$ para todo a de E. Demostrar que f es suprayectiva y estrictamente creciente. ¿Qué se puede concluir?

29. Sea $k \to a_k$ una aplicación estrictamente creciente de **N** en **N** tal que $a_0 = 0$.

a) Demostrar que el conjunto (a_k) $(k \in \mathbf{N})$ no está acotado superiormente.

b) Demostrar que para todo entero natural x, existe un entero natural k único tal que

$$a_k \leq x < a_{k+1}.$$

30. Siendo a un entero natural superior o igual a 2, demostrar que:

a) Para todo entero natural $x \neq 0$, existe un entero natural único k tal que

$$a^k \leq x < a^{k+1}.$$

b) Para todo entero $x \neq 0$, existe un entero natural único k y un entero natural x_k único verificando $0 \leq x_k < a$ tales que

$$x_k a^k \leq x < (x_k + 1)a^k.$$

c) Para todo entero natural x existe un entero natural n único y una sucesión única de enteros naturales $x_n, x_{n-1}, \ldots, x_1, x_0$ verificando $0 \leq x_k < a$ para todo k de $[0, n]$ tal que

$$x = x_n a^n + \ldots + x_1 a + x_0.$$

---

* Obsérvese que este ejercicio aparece resuelto en el § 28.

31. Demostrar que 1000! es divisible por 2^{994} y no por 2^{995}. Buscar el mayor entero n tal que 3^n divide 1000!

32. Encontrar el número de soluciones enteras y positivas o nula de la ecuación

$$x + 2y = n$$

(dado n entero natural).

33. Siendo n un entero natural, sea u_n el número de sucesiones finitas $(k_1, k_2, \dots, k_p)$, donde cada k_i es igual a 1 o a 2, tales que

$$k_1 + k_2 + \dots + k_p = n + 1.$$

Demostrar que

$$u_{n+2} = u_{n+1} + u_n.$$

(Sea S_n el conjunto de las sucesiones, se demostrará que el conjunto S'_n de las sucesiones terminadas por 1 y el conjunto S''_n de sucesiones terminadas por 2 son complementarios.)

34. a) Demostrar que para $0 \leq k \leq p \leq n$

$$C_n^k C_{n-k}^{p-k} = C_n^p C_p^k.$$

([illegible] n elementos, se podrá considerar entre las partes con p elementos aquellas [illegible] una parte determinada que tiene k elementos.)

[illegible]

$$C_n^0 C_n^p + C_n^1 C_{n-1}^{p-1} + \dots + C_n^p C_{n-p}^0 = 2^p C_n^p.$$

35. Siendo E un conjunto finito de n elementos, se designa por $\mathfrak{F}$ el conjunto de los recubrimientos de E, (X_i) $(1 \leq i \leq m)$ por partes disjuntas dos a dos tales que card $X_i = p_i$ (con $p_1 + \dots + p_m = n$). Sea (A_i) y (B_i) $(1 \leq i \leq m)$ dos recubrimientos de E que cumplen esta condición. Demostrar que el número de permutaciones f de E tales que $f(A_i) = B_i$ para todo i de $[1, m]$ es $p_1!\, p_2!\, \dots, p_m!$ (Se considera la aplicación f_i de A_i en B_i que coincide con f sobre A_i; ver § 40, ej. 2.)

36. Las mismas notaciones que en el ejercicio 35, siendo f una permutación de E, demostrar que si (A_i) $(1 \leq i \leq m)$ es un elemento de $\mathfrak{F}$, análogamente lo es $(f(A_i))$ $(1 \leq i \leq m)$. Escogido (A_i) $(1 \leq i \leq m)$ y siendo P el conjunto de las permutaciones de E, se considera la aplicación $g\colon P \to \mathfrak{F}$ definida por

$$g(f) = (f(A_i)) \ (1 \leq i \leq m).$$

Demostrar que g es suprayectiva.

Si (X_i) $(1 \leq i \leq m)$ es un elemento de $\mathfrak{F}$, ¿cuál es el cardinal $g^{-1}(X_i)$?

Deducir el cardinal de $\mathfrak{F}$. (Utilizar el ejercicio precedente y el principio de los pastores). ¿Qué relación tiene este ejercicio con el cálculo de C_n^p (§ 41)?

37. Dado un conjunto E con n elementos, se llaman *combinaciones de estos n elementos tomados m a m con repetición* todo conjunto con m elementos de E, pudiéndose repetir cada uno de ellos hasta m veces.

(Por ejemplo, si $E = \{a, b, c, d, e, f\}$, $\{a, a, c, d, d, f\}$ es una combinación con repeticiones de 6 elementos a, b, c, d, e, f, tomados 7 a 7.)

Se designa por Γ_n^m el número de estas combinaciones con repetición:

a) Demostrar que un elemento determinado figura $\frac{m}{n}\Gamma_n^m$ veces en el conjunto de las combinaciones con repetición de n elementos tomados m a m.

b) Demostrar que

$$\frac{m}{n}\Gamma_n^m = \Gamma_n^{m-1} + \frac{m-1}{n}\Gamma_n^{m-1}$$

deducir que

$$\Gamma_n^m = C_{m+n-1}^m.$$

c) Demostrar que

$$\Gamma_n^0 + \Gamma_n^1 + \dots + \Gamma_n^m = \Gamma_{n+1}^m.$$

38. En el plano referido a dos ejes rectangulares de vectores unitarios respectivos $\vec{i}$ y $\vec{j}$, se considera las líneas quebradas uniendo O(0, 0) al punto M(p, q) (p y q enteros naturales), cuyos lados son equipotentes, bien a $\vec{i}$, bien a $\vec{j}$, sea b el número de líneas quebradas.

a) Demostrar que cada una de estas líneas quebradas tiene $p + q$ lados que se numerarán $1, 2, \dots, p + q$. Demostrar que el conocimiento del número de lados equipotentes a $\vec{i}$ (o a $\vec{j}$) determina la línea quebrada. Deducir que

$$b = C_{p+q}^p = C_{p+q}^q.$$

b) Demostrar que los lados equipotentes a $\vec{i}$ (respectivamente, a $\vec{j}$) tienen $q +$ ordenadas (respectivamente, $p + 1$ abscisas) posibles (ver ejercicio precedente para la definición de Γ_n^m) $b = \Gamma_{q+1}^p = \Gamma_{p+1}^q$.

c) Demostrar que $\Gamma_n^m = C_{m+n-1}^m$.

39. a) Siendo a y n dos enteros naturales no nulos, demostrar por recurrencia que

$$(1 + a)^n \geq 1 + na.$$

b) Demostrar que si $n > 2$ y $p > 1$ la igualdad $n^{p-1} = p$ es imposible; para $n =$ demostrar que la igualdad precedente es posible sólo para $p = 1$ o $p = 2$.

c) Demostrar que la igualdad (m, n, p enteros naturales no nulos)

$$(m^n)^p = m^{(n)^p}$$

es exactamente únicamente para las ternas $(1, n, p)$, $(m, n, 1)$, $(m, 2, 2)$.

40. Si x e y son dos enteros naturales no nulos tales que $x < y$, se propone resolver la ecuación

$$x^y = y^x.$$

a) Sea Δ el m.c.d. de x e y, se pone $x = x'\Delta$, $y = y'\Delta$.
Demostrar que x' divide y'^{x-1} y por recurrencia que divide y'. Deducir que $\Delta = x$.
b) Utilizando el ejercicio precedente, demostrar que $x = 2$, $y = 4$.

[illegible]. Se considera la aplicación de $\mathbf{N} \times \mathbf{N}$ en $\mathbf{N}$ definida por

$$\begin{array}{lll} x \geq 1 & & f(x, y) = f(x-1, y+1) + 1 \\ x = 0 & y \geq 1 & f(0, y) = f(y-1, 0) + 1 \\ & & f(0, 0) = 0. \end{array}$$

Siendo k un entero natural, se llamará segmento s_k el conjunto de las parejas de enteros naturales (x, y) tales que $x + y = k$.

a) ¿Cuál es el número de elementos de s_k.
Calcular $f(x, y)$ en función de x y de y.

b) Siendo n un entero natural, demostrar que existe un entero natural k único tal que

$$\frac{k(k+1)}{2} \leq n < \frac{(k+1)(k+2)}{2}$$

[illegible] utilizar el ejercicio 20 y las inecuaciones de segundo grado).
[illegible] que cualquiera que sea n la ecuación

$$f(x, y) = n$$

[illegible] única (se tratará de determinar el orden k del segmento s_k, al [illegible] (x, y)). Deducir que f es biyectiva.
[illegible] calcular x e y para $n = 1970$.

3

LEYES DE COMPOSICION

43. En el capítulo I hemos estudiado las propiedades generales de los conjuntos que sólo utilizan las nociones de pertenencia y de orden; en este capítulo vamos a definir y estudiar las *operaciones algebraicas* relativas a uno o varios conjuntos: generalizan las operaciones elementales definidas sobre **N** en el capítulo 2; esta operación —o estas operaciones— definida sobre un conjunto E, le da una *estructura algebraica* de una especie determinada, que depende de la operación —o de las operaciones— considerada; el conjunto E se llama el *soporte* de la estructura considerada: *un mismo conjunto* E *puede estar provisto de estructuras diferentes, y recibe entonces calificaciones diferentes* (ver § 68).

El *Algebra* es la rama de las matemáticas que estudia sistemáticamente las estructuras algebraicas: es decir, que se interesa con prioridad por las *propiedades de las operaciones* definidas entre los elementos de los diversos conjuntos que por los mismos *elementos*. Naturalmente, definiremos los conjuntos clásicos **Z**, **Q**, **R**, **C**, polinomios; volveremos a encontrar sus propiedades y definiremos otras.

En este capítulo vamos a definir ciertas *operaciones algebraicas* (leyes de composición interna y leyes de composición externa) y estudiar sus propiedades generales; el resto del libro se dedicará al estudio de las *estructuras fundamentales del Algebra.*

1. Conjuntos provistos de una ley de composición interna

44. Definición de una ley de composición interna. Notaciones. Ejemplos

DEFINICIÓN. — *Se llama* ley de composición interna *entre elementos de* E, *toda aplicación de una parte* A *de* E × E *en* E. *Se dice entonces que* E *está* provisto de la ley *interna considerada.*

Es decir, a la pareja (x, y) de A, una ley interna hace corresponder un elemento único z de E

$$(x, y) \to z = f(x, y).$$

Cuando A = E × E, se dice que *la ley está definida sobre todo* E, se dice entonces que es una *operación interna sobre* E o una *ley interna sobre* E.

En lo sucesivo, salvo mención contraria, las leyes estudiadas se supondrán definidas sobre todo E.

El valor $f(x, y)$ que corresponde a la pareja (x, y) se llama el *compuesto* de x y de y que son sus *términos*. El compuesto se designa con diversos símbolos:

$$x + y, \quad x \cdot y, \quad xy, \quad x \top y, \quad x \perp y, \quad x \Omega y, \quad x * y$$

símbolos específicos de algunas leyes (ver los ejemplos).

La operación $(x, y) \to x + y$ se llama *adición*; $x + y$ (que se lee "*x más y*") es la *suma* de los dos términos x e y.

La operación $(x, y) \to x \cdot y$ o $(x, y) \to xy$ se llama *multiplicación*; $x \cdot y$ o xy —que se lee "*x multiplicado por y*" o simplemente "*xy*"— es el *producto* de los dos *factores* x e y.

En estos dos casos se dice que se ha adoptado, respectivamente, la *notación aditiva* y la *notación multiplicativa* para la ley considerada.

$x \top y$ se lee algunas veces "*x truc y*" y $x \perp y$ "*x antitruc y*".

Dada la ley interna $(x, y) \to x \top y$, la ley interna $(x, y) \to y \top x$ se llama *opuesta a la primera*.

Observemos que siendo único el compuesto de dos términos cualesquiera que sean x, y, z es

$$x = y \Rightarrow x \top z = y \top z \quad \text{y} \quad z \top x = z \top y.$$

En fin, *por abuso de lenguaje*, en lugar de decir "consideremos sobre E la ley $(x, y) \to x \top y$", diremos "consideremos la ley $\top$ sobre E" o también "consideremos la ley $x \top y$ definida sobre E".

En resumen, podemos decir que *un conjunto* E *provisto de una ley interna* $\top$ *es un par, lo que expresaremos frecuentemente por* (E, $\top$).

Así, (N, +) y (N, ×) designarán, respectivamente, el conjunto de los enteros naturales provisto de la adición y de la multiplicación.

EJEMPLOS

1. Las dos leyes $a + b$, ab son leyes internas definidas sobre todo **N**, no ocurre lo mismo con las leyes $a - b$ (definida solamente para $b \leq a$) y a/b (definida solamente para $b \mid a$).

2. En **N** la ley a^b está definida para toda pareja (a, b), salvo $(0, 0)$.

3. $a + b$, $a - b$, ab son leyes definidas sobre todo **Z**, **Q**, **R**, **C**. No ocurre lo mismo con a/b, que sólo está definida sobre **Z** cuando b divide a a y sobre **Q**, **R**, **C** cuando $b \neq 0$.

4. Siendo E un conjunto, las leyes $A \cup B$, $A \cap B$ están definidas sobre todo el conjunto $\mathfrak{F}(E)$.

5. Siendo E un conjunto, $f \circ g$ es una ley interna definida sobre todo el conjunto $\mathfrak{F}(E, E)$ (ver § 15).

Observemos que si E, F, G son tres conjuntos, f un elemento de $\mathfrak{F}(E, F)$ y g un elemento de $\mathfrak{F}(F, G)$, $(f, g) \to g \circ f$ es una aplicación de $\mathfrak{F}(E, F) \times \mathfrak{F}(F, G)$ en $\mathfrak{F}(E, G)$, que no es en general una ley interna.

6. Las leyes internas sup (a, b), inf (a, b) están definidas sobre la totalidad de todo conjunto E totalmente ordenado. Si el conjunto E está parcialmente ordenado, estas dos leyes no están en general totalmente definidas sobre el conjunto, pero sí lo son si E es un retículo por la relación de orden considerada (ver capítulo 1, ej. 24).

7. En $\mathbf{R}^3$, el producto vectorial $\vec{A} \wedge \vec{B}$ de dos vectores es una ley interna (no ocurre lo mismo con el producto escalar, pues $\vec{A} \cdot \vec{B}$ es un número real y no un vector).

Cuando E es *finito*, la ley está determinada por su *tabla* (análoga a la tabla de PITÁGORAS); por ejemplo, en notación multiplicativa

← factor de la derecha

	a_1	…	a_p	…	a_q	…	a_n
a_1	$(a_1)^2$	…	a_1a_p	…	a_1a_q	…	a_1a_n
.	.		.		.		.
.	.		.		.		.
.	.		.		.		.
a_p	a_pa_1	…	$(a_p)^2$	…	a_pa_q	…	a_pa_n
.	.		.		.		.
.	.		.		.		.
.	.		.		.		.
a_q	a_qa_1	…	a_qa_p	…	$(a_q)^2$	…	a_qa_n
.	.		.		.		.
.	.		.		.		.
.	.		.		.		.
a_n	a_na_1	…	a_na_p	…	a_na_q	…	$(a_n)^2$

↑ factor de la izquierda

OBSERVACION

En la notación ab o $a \top b$, es mejor designar a como el «*factor —o término— de izquierda*» en lugar de «primer factor —o término—», pues en la notación $g \circ f$, por ejemplo, podría surgir ambigüedad: f «factor de derecha» se efectúa en primer lugar y g «factor de izquierda» se efectúa en segundo (ver § 15).

[illegible] Composición de una sucesión ordenada finita de elementos. Asociatividad

Sea tres elementos a, b, c dados en este orden de (E, T), para definir el elemento compuesto de a, b, c en este orden, se puede calcular los elementos compuestos

$$(a \top b) \top c \qquad \text{o} \qquad a \top (b \top c)$$

Puede ocurrir que estos resultados sean desiguales; como vamos a ver, los cálculos en (E, T) son mucho más fáciles si estos dos elementos compuestos son iguales, cualesquiera que sean a, b, c; de donde la definición siguiente:

Definición. *Una ley interna T definida sobre* E *es* asociativa *si y sólo si*

$$(\forall a, b, c \in E) \qquad (a \top b) \top c = a \top (b \top c).$$

Observemos (ver § [illegible], nota [illegible]) que la notación $(\forall a, b, c \in E)$ es un abuso de notación por $(\forall (a, b, c) \in E \times E \times E)$ o $(\forall a \in E)(\forall b \in E)(\forall c \in E)$.

Así, sobre un conjunto provisto de una ley asociativa escrita multiplicativamente se tiene

$$((ab)c)d = (a(bc))d = a((bc)d) = a(b(cd)),$$

[illegible] se puede empezar el cálculo, ya por la izquierda, ya por la [illegible]

[illegible] igualmente para una ley asociativa el elemento compuesto de [illegible] a_n dados en este orden por

$$\top a_1 = a_1$$

$$1 \leq p \leq n \quad \mathop{\top}_{i \in [1, p]} a_i = \left(\mathop{\top}_{i \in [1, p-1]} a_i \right) \top a_p$$

[illegible] el cálculo por la izquierda, las demostraciones fáciles en su principio pero pesadas, que omitiremos, muestran que el resultado sería el mismo si se hubiera empezado el cálculo por la derecha y más generalmente *si los paréntesis estuvieran colocados de manera arbitraria, respetándose el orden de los diferentes términos* a_1, a_2, ..., a_n; proponemos estas demostraciones como ejercicio (ver fin del capítulo, ej. 42).

Este resultado único para una ley asociativa será expresado sin paréntesis, según la notación

$$a_1 \top a_2 \dots \top a_n = \mathop{\top}_{i=1}^{i=n} a_i$$

$$a_1 + a_2 \dots + a_n = \sum_{i=1}^{i=n} a_i$$

$$a_1 \quad a_2 \dots \quad a_n = \prod_{i=1}^{i=n} a_i$$

la observación hecha en el § 32 (ver también § 6, *b*, nota 2) sobre los índices es válida aquí: el índice i puede reemplazarse por un símbolo cualquiera

salvo 1 (*su primer valor*) y n (*su último valor*).

El elemento compuesto de una sucesión de n elementos iguales a a ($n \geq 1$) se designará, respectivamente,

$$\mathop{\top}_{1}^{n} a, \qquad na, \qquad a^n$$

conviniendo es

$$\mathop{\top}^{1} a = a, \qquad 1a = a, \qquad a^1 = a.$$

En la última notación se dice que a^n es la *potencia* de *exponente* n de a.

Se verificará que para todo entero $p \geq 1$ y todo entero $q \geq 1$, se tiene

$$\begin{cases} \left(\mathop{\top}^{p} a\right)\left(\mathop{\top}^{q} a\right) = \mathop{\top}^{p+q} a \\ pa + qa = (p+q)a \\ a^p a^q = a^{p+q} \end{cases}$$

$$\begin{cases} \mathop{\top}^{p} \left(\mathop{\top}^{q} a\right) = \mathop{\top}^{pq} a \\ p(qa) = (pq)a \\ (a^p)^q = a^{pq}. \end{cases}$$

EJEMPLOS

1. Las leyes $a + b$, ab en **N**, **Z**, **Q**, **R**, **C** son asociativas, análogamente las leyes definidas en los ejemplos 4, 5, 6 del párrafo precedente.

2. La ley a^b definida en **N**, salvo para (0, 0), no es asociativa (ver § 37), análogamente para el producto vectorial (ejemplo 7).

3. La ley $f \circ g$ en $\mathfrak{F}$(E, E) (§ 44, ejemplo 5) es asociativa.

Para una ley *no asociativa* nunca se empleará la notación $\mathop{\top}\limits_{i=1}^{i=n} a_i$, se indicará siempre cómo está efectuada la operación (ver § 37 para la exponenciación en **N**).

Salvo mención contraria, las leyes internas utilizadas en lo sucesivo se considerarán asociativas.

EJERCICIOS

1. En **R**, dados dos números reales a y b, mostrar que entre las leyes

$$x \perp y = ax + by$$

una sola es asociativa.

3. Estando E provisto de una ley asociativa T, a un elemento fijo de E, mostrar que la ley ⊥ definida por

$$x \perp y = x \top a \top y$$

es asociativa.

[illegible]. Conmutatividad de una ley interna. Elementos permutables

a) Definición. — *Una ley interna* T *definida sobre* E *es conmutativa si y sólo si*

$$(\forall a, b \in E) \qquad a \top b = b \top a.$$

EJEMPLOS

1. Las leyes $a + b$ y ab en N, Z, Q, R, C son conmutativas, análogamente las leyes definidas en los ejemplos 4 y 6 del § 44.

2. Las leyes definidas en los ejemplos 2, 5, 7 del § 44 no son conmutativas.

[illegible] una ley T definida sobre E a la vez *asociativa y conmutativa*; se de[illegible] (ver [illegible], 44, [illegible] del capítulo) que [illegible] $i \to j(i)$ es una [illegible] sobre sí mismo

$$\mathop{\top}_{i=1}^{i=n} a_i = \mathop{\top}_{i=1}^{i=n} a_{j(i)}$$

[illegible] más general se demostrará que si $I_1, I_2, \ldots, I_p$ es una [illegible] un conjunto finito I de índices; se tiene

$$\mathop{\top}_{i \in I} a_i = \mathop{\top}_{k=1}^{k=p} \left(\mathop{\top}_{i_k \in I_k} a_{i_k} \right).$$

El lector escribirá estas fórmulas en forma aditiva y multiplicativa.

Por ejemplo, si el conjunto de índices es $K = I \times J$, con $I = [1, p]$ y $J = [1, q]$, los elementos a_{ij} con estos índices (ver § 17) pueden escribirse en forma de una tabla rectangular (matriz)

$$\begin{matrix} a_{11} & a_{21} & \ldots & a_{i1} & \ldots & a_{p1} \\ a_{12} & a_{22} & \ldots & a_{i2} & \ldots & a_{p2} \\ \cdot \\ \cdot \\ \cdot \\ a_{1j} & a_{2j} & \ldots & a_{ij} & \ldots & a_{pj} \\ \cdot \\ \cdot \\ \cdot \\ a_{1q} & a_{2q} & \ldots & a_{iq} & \ldots & a_{pq}. \end{matrix}$$

Los elementos de una partición de K pueden ser o bien los índices d los elementos de una misma columna (i constante), o bien los índices de l elementos de una misma línea (j constante); como es la notación aditiva qu se empleará en álgebra lineal, damos la fórmula relativa a la suma de tod los elementos de la tabla, por

$$(3) \qquad \sum_{i=1}^{i=p}\left(\sum_{j=1}^{j=p} a_{ij}\right)=\sum_{j=1}^{j=q}\left(\sum_{i=1}^{i=p} a_{ij}\right)=\sum_{i=1}^{i=p}\sum_{j=1}^{j=q}(a_{ij}).$$

El símbolo utilizado en el último miembro de (3) está definido por l fórmula precedente, o aplicando la fórmula (2) (en notación aditiva) por

$$\sum_{i=1}^{i=p}\sum_{j=1}^{j=q} a_{ij}=\sum_{(i,\,j)\in \mathrm{I}\times\mathrm{J}} a_{ij}.$$

En lo sucesivo:

Convendremos que una ley representada aditivamente es asociativa conmutativa.

b) Sin que la ley T sea conmutativa, puede suceder que para dos el mentos particulares a y b de E, se tenga

$$a \top b = b \top a$$

se dice entonces que a y b son *permutables* en la ley T, o también que a y *permutan* en la ley T. Se dice también que a y b son *conmutables* o tambi que *conmutan* en la ley T.

Si existe un elemento a permutable con todo elemento x de E, es deci tal que

$$(\forall\, x \in \mathrm{E}) \qquad a \top x = x \top a$$

se dice que a es un *elemento central* de (E, T). Se llama *centro* de (E, T) conjunto de sus elementos centrales.

EJERCICIOS

1. Si a y b son dos elementos permutables de (E, T), mostrar que ocurre lo mis para $\mathop{\top}\limits^{p} a$ y $\mathop{\top}\limits^{q} b$ y que

$$\mathop{\top}\limits^{n}(ab)=\left(\mathop{\top}\limits^{n} a\right)\top\left(\mathop{\top}\limits^{n} b\right).$$

2. Las dos proposiciones (a y b elementos de (E, T))

$$(\exists\, a)(\forall\, b) \qquad a \top b = b \top a$$
$$(\forall\, b)(\exists\, a) \qquad a \top b = b \top a$$

¿son equivalentes? Interpretarlas.

3. Buscar el error en el razonamiento siguiente: «ninguna ley no puede ser no conmutativa; en efecto, la relación

$$a \top b \neq b \top a$$

conduciría, para $a = b$, a una contradicción».

4. Demostrar que en $\mathcal{F}(E, E)$, la ley $f \circ g$, no es conmutativa si E tiene al menos dos elementos, pero que id_E permuta con toda aplicación.

17. Elementos regulares

Estando E provisto de una ley $\top$, consideremos las dos aplicaciones de E en E definidas por

$$x \to \gamma_a(x) = a \top x$$
$$x \to \delta_a(x) = x \top a$$

Se llaman, respectivamente, *traslación por la izquierda* y *traslación por la derecha* asociadas al elemento a.

[illegible] inyectiva, para todo x y todo y de E

$$a \top x = a \top y \rightarrow x = y$$

[illegible] que a es *regular por la izquierda*.

[illegible] si δ_a es *inyectiva*, para todo x y todo y de E

$$x \top a = y \top a \rightarrow x = y$$

[illegible] que a es *regular por la derecha*.

Un elemento a es *regular* si es a la vez regular por la izquierda y por la derecha; cuando se pasa de la igualdad $a \top x = a \top y$ (o $x \top a = y \top a$) a la [illegible], se dice que se *simplifica por* a.

EJERCICIOS

1. Demostrar que para la adición en **R** (o **Q**, **Z**, **N**) todo elemento es regular. ¿Ocurre lo mismo para la multiplicación?

2. ¿Cuáles son los elementos regulares en $\mathcal{S}(E)$ para la reunión?, ¿para la intersección?

3. Demostrar que $\mathcal{F}(E, E)$ provisto de la ley $f \circ g$, los elementos regulares por la izquierda son las inyecciones y los elementos regulares por la derecha las suprayecciones (ver § 10, ej. 2).

4. Si E está provisto de la ley asociativa $\top$, demostrar que

$$\gamma_{a \top b} = \gamma_a \circ \gamma_b \qquad \delta_{a \top b} = \delta_b \circ \delta_a .$$

48. Elemento neutro. Elementos simetrizables. Elementos simétricos

a) Teorema y definición. — *Si existe en* (E, T) *un elemento e tal que*

$$(\forall x \in E) \qquad e \top x = x \top e$$

si la ley T *es asociativa, es único, se le llama* el elemento neutro *de* (E, T).

Sea, en efecto, dos elementos e y e' tales que, para todo x de E,

$$(1) \quad e \top x = x \top e = x \qquad (2) \quad e' \top x = x \top e' = x$$

haciendo $x = e'$ en (1) y $x = e$ en (2), se tiene

$$e \top e' = e' \top e = e' \qquad e' \top e = e \top e' = e.$$

Se ve que el elemento neutro, si existe, es un elemento central de (E, T). En notación *aditiva* el elemento neutro se llama *elemento cero*; en notación *multiplicativa, elemento unidad.*

EJEMPLOS Y EJERCICIOS

1. En **R** (o **Q**, **Z**, **N**) 0 es el elemento neutro de la adición y 1 el elemento neutro de la multiplicación. El conjunto de los enteros racionales pares (o múltiplos de $a \neq 1$) ¿tiene un elemento neutro para la multiplicación?
2. ¿Cuál es el elemento neutro de $\mathfrak{P}$(E) para la reunión?, ¿para la intersección?
3. ¿Cuál es el elemento neutro de $\mathfrak{F}$(E, E) para la ley $f \circ g$?
4. En **N*** las leyes que a (a, b) hacen corresponder el m. c. d., o el m. c. m., de a y b, ¿tienen un elemento neutro?

b) Definición. — *Dada una ley* T *definida sobre* E *poseyendo un elemento neutro e, un elemento a es* simetrizable *por la ley* T *si existe a' de* E *tal que*

$$a \top a' = a' \top a = e.$$

Se dice que a' es un simétrico *de a en* (E, T).

Siendo la relación precedente simétrica en a y a': a es un simétrico de a', se dice que a y a' son *simétricos*.

Teorema. — *Si para una ley* T *definida sobre* E, *asociativa, poseyendo un elemento neutro e, a es simetrizable, su simétrico es único y a es regular.*

Sean a' y a'' dos simétricos de a; tenemos

$$(1) \quad a \top a' = a' \top a = e \qquad (2) \quad a \top a'' = a'' \top a = e.$$

Utilicemos estas dos relaciones para calcular de dos maneras, merced a la asociatividad, $a'' \top a \top a'$,

$$a'' \top a \top a' = (a'' \top a) \top a' = e \top a' = a'$$
$$a'' \top a \top a' = a'' \top (a \top a') = a'' \top e = a'',$$

luego $a' = a''$. Demostremos ahora que todo elemento simetrizable es regular.

Compongamos por la izquierda (o por la derecha) los dos miembros de $a \top x = a \top y$ (o $x \top a = y \top a$) con a', obtendremos

$$\begin{aligned} a \top x &= a \top y \\ a' \top (a \top x) &= a' \top (a \top y) \\ (a' \top a) \top x &= (a' \top a) \top y \\ e \top x &= e \top y \\ x &= y \end{aligned} \qquad \begin{aligned} x \top a &= y \top a \\ (x \top a) \top a' &= (y \top a) \top a' \\ x \top (a \top a') &= y \top (a \top a') \\ x \top e &= y \top e \\ x &= y, \end{aligned}$$

luego a es regular.

TEOREMA. — *Para toda ley* $\top$ *asociativa, poseyendo un elemento neutro e, el elemento compuesto de dos elementos simetrizables es simetrizable y se tiene*

$$(a \top b)' = b' \top a'.$$

Sean a' y b' los simétricos respectivos de los elementos simetrizables a y b, supongamos que existe c tal que

$$(1) \quad c \top (a \top b) = e \qquad (2) \quad (a \top b) \top c = e$$

compongamos (1) sucesivamente con b' y a'; por la derecha obtendremos

$$[c \top (a \top b)] \top b' = (c \top a) \top (b \top b') = (c \top a) \top e = c \top a = e \top b' = b'$$

y como

$$(c \top a) \top a' = c \top (a \top a') = c \top e = c = b' \top a'$$

se ve fácilmente que $c = b' \top a'$ verifica también (2); luego $a \top b$ es simetrizable y $b' \top a'$ es su simétrico.

EJERCICIO

Demostrar que, en E provisto de una ley asociativa poseyendo un elemento neutro, si a, que admite un simétrico a', es permutable con b, entonces a' es también permutable con b; deducir de lo anterior que el simétrico de todo elemento central simetrizable es un elemento central.

12. Observaciones sobre las ecuaciones $\gamma_b(x) = a$, $\delta_b(x) = a$. Notaciones diversas

Supongamos E provisto de una ley asociativa $\top$ poseyendo un elemento neutro e, consideremos las ecuaciones

$$(1) \quad \gamma_b(x) = b \top x = a \qquad (2) \quad \delta_b(x) = x \top b = a.$$

Si b tiene un simétrico b', componiendo por la izquierda, o por la derecha con b', se obtiene por solución respectiva de (1) y (2)

$$(1') \quad x_1 = b' \top a \qquad (2') \quad x_2 = a \top b';$$

estas soluciones son, en general, distintas para una ley no conmutativa.

En notación aditiva, puesto que la ley es entonces conmutativa, las ecuaciones (1) y (2) se reducen a

$$b + x = a.$$

La solución, si existe, $x = a + b'$ expresada $x = a - b$, se llama *diferencia* de a y b. Si se hace $a = e$ se ve que el simétrico de b está representado por $(-b)$, que se llama entonces *opuesto* de b. Luego se tiene

$$a - b = a + (-b)$$

se verificará que

$$-(a - b) = b - a$$

En notación multiplicativa el simétrico de b se llama *inverso* de b (se dice que b es *inversible*) y se designa b^{-1}, las soluciones (si existen) de

$$bx = a \qquad xb = a$$

son, respectivamente,

$$x_1 = b^{-1}a \qquad x_2 = ab^{-1}$$

y se llaman, respectivamente, *cociente por la izquierda* y *cociente por la derecha* de a por b.

Cuando la *ley es conmutativa* se escribe

$$x = x_1 = x_2 = a/b$$

x es el *cociente* de a por b, a/b es una *fracción de la cual* a es el *numerador* y b el *denominador.*

II. Diversas leyes internas asociadas a una misma ley interna

50. Ley interna definida sobre $\mathfrak{P}$(E) deducida de una ley interna definida sobre E

Estando E provisto de una ley interna T, podemos definir una ley interna T′ sobre $\mathfrak{P}$(E) poniendo, para A y B dos partes de E,

$$\mathrm{A\,T'\,B} = \{z \mid (\exists\, x \in \mathrm{A}),\ \ (\exists\, y \in \mathrm{B}),\ \ z = x\,\mathrm{T}\,y\},$$

dicho de otra manera, A T′ B es la parte de E descrita por $x \mathrm{T} y$ cuando x describe A e y describe B.

Las leyes T y T′ están definidas sobre dos conjuntos diferentes E y $\mathfrak{P}$(E); por otra parte,

$$\{x\}\,\mathrm{T'}\,\{y\} = \{x\,\mathrm{T}\,y\}.$$

En general no hay ningún inconveniente en designar las dos leyes por el mismo símbolo T, se dice entonces que la ley

$$(A, B) \rightarrow A \top B$$

definida sobre $\mathfrak{P}(E)$ es la *extensión de la ley* $x \top y$ *a las partes* de E. En ciertos casos no se puede representar las dos leyes por el mismo símbolo (ver ejercicio 3 más abajo).

En notación aditiva, o multiplicativa, el elemento compuesto de dos partes A, B estará expresado A + B y AB. En este último caso, se procurará no confundir $A^2 = AA$ parte de E provista de una multiplicación, y $A^2 = A \times A$ parte del producto cartesiano E × E.

Por abuso de notación $\{x\} \top A$, $\{x\} + A$, $\{x\}A$ se expresarán $x \top A$, $x + A$, xA. Por ejemplo, 2N designará el conjunto de los enteros naturales pares, aN el conjunto de los enteros naturales múltiples del entero natural a (ver § 34, *c*).

Cuando hayamos definido el conjunto **Z** de los enteros racionales, —**N** designará el conjunto de los enteros negativos o nulos, a**Z** el conjunto de los enteros múltiplos de a.

EJERCICIOS

1. Cuando la ley no está definida sobre todo E, si se conviene en poner $\{x\} \top \{y\} = \varnothing$ cuando $x \top y$ no está definido, demostrar que la extensión a las partes está completamente definida sobre $\mathfrak{P}(E)$. Determinar en $\mathfrak{P}(N)$, A — B cuando

$$A = [3, 7] \quad B = [5, 11] \quad \text{o} \quad A = [a, a + h] \quad B = [b, b + k].$$

2. Demostrar que si la ley en E es asociativa o conmutativa, también lo es su extensión a las partes de E. Si la ley en E tiene un elemento neutro, ¿también lo tiene la extensión a las partes?

3. Consideremos la ley A ∪ B definida sobre $\mathfrak{P}(E)$, ¿se puede designar $\mathfrak{A} \cup \mathfrak{B}$, la extensión a las partes que aquí es una ley definida sobre $\mathfrak{P}(\mathfrak{P}(E))$?

4. Sea E un conjunto provisto de una ley asociativa poseyendo un elemento neutro e, representado multiplicativamente, y una parte A descrita por los elementos inversibles de E. Comparar AA^{-1} y $\{e\}$.

5. Sea E un conjunto provisto de una ley representada multiplicativamente, demostrar que si a es regular, $\{a\}$ es regular en $\mathfrak{P}(E)$. Siendo A una parte de E descrita por elementos regulares, demostrar mediante el ejemplo

$$N^* \{2, 3\} = N^* \{2, 3, 6\}$$

que A no es forzosamente regular en $\mathfrak{P}(E)$.

41. Parte estable. Ley inducida

DEFINICIÓN. — *Una parte* A *de un conjunto* E *provisto de una ley interna* T *es* **estable** *para esta ley si*

$$(\forall x, y \in A) \qquad x \top y \in A.$$

Utilizando las notaciones del párrafo precedente se ve, pues, que A es estable si y sólo si $A \top A \subset A$.

La aplicación de $A \times A$ en A definida por

$$(x, y) \to x \top y$$

es, pues, una ley de composición interna definida sobre A, se llama *ley inducida sobre* A *por la ley* T *definida sobre* E, se dice que la ley T sobre E *prolonga* la ley así definida sobre A. Si no da lugar a confusión, la ley T definida sobre E y la ley inducida por T sobre una parte estable A de E se designarán con el *mismo símbolo.*

Si la ley T es asociativa o conmutativa sobre E, se ve inmediatamente que la ley inducida en una parte estable de E es asociativa y conmutativa.

Se ve igualmente que si la ley T posee un elemento neutro *e* perteneciente a una parte estable A, la ley inducida sobre A admite *e* por elemento neutro, pero éste no es el único caso a tratar (ver ejercicios 1 y 2 más abajo).

EJERCICIOS

1. $2N^*$ es una parte estable de N^* para la multiplicación, la ley inducida está desprovista de elemento neutro.

2. Sea E un conjunto, A una parte propia de E; se designa por $\mathcal{A}_1$ el conjunto de las partes X_1 de E tales que $A \subset X_1$ y por $\mathcal{A}_2$ el conjunto de las partes X_2 de E tales que $X_2 \subset A$; demostrar que $\mathcal{A}_1$ y $\mathcal{A}_2$ son partes estables de $\mathcal{P}(E)$ para la reunión y la intersección. Los tres conjuntos $\mathcal{P}(E)$, $\mathcal{A}_1$ y $\mathcal{A}_2$ tienen cada uno un elemento neutro para la reunión y las leyes inducidas, ¿son las mismas? La misma cuestión para la intersección.

3. Demostrar que el conjunto de los elementos regulares de E para una ley asociativa definida sobre E es estable para esta ley.

4. Todo elemento regular *a* para una ley definida sobre E es regular para la ley inducida sobre una parte estable de E conteniendo *a*. ¿Es verdadera la recíproca?

52. Ley producto

Dados dos conjuntos E_1 y E_2 provistos, respectivamente, de una ley interna $\top_1$ y de una ley interna $\top_2$, se puede definir una ley T sobre el producto cartesiano $E_1 \times E_2$ poniendo para dos parejas cualesquiera (a_1, a_2) y (b_1, b_2)

$$(a_1, a_2) \top (b_1, b_2) = (a_1 \top_1 b_1, a_2 \top_2 b_2).$$

La ley T se llama ley producto de las leyes $\top_1$ y $\top_2$. Se ve fácilmente que si $\top_1$ y $\top_2$ son asociativas o conmutativas, también lo es la ley T. Si $\top_1$ y $\top_2$ tienen por elementos neutros respectivos e_1 y e_2, la ley T tiene un elemento neutro $e = (e_1, e_2)$.

Si no se teme ninguna confusión se podrá designar las tres leyes $\top_1$, $\top_2$, T por un mismo símbolo. Por ejemplo, si $E_1 = E_2 = E$ sobre $E \times E$, podremos definir una ley producto y tendremos

$$(a, a') \top (b, b') = (a \top b, a' \top b')$$
$$(a, a') + (b, b') = (a + b, a' + b')$$
$$(a, a')(b, b') = (ab, a'b')$$

según la notación utilizada.

Se puede generalizar esto a $E = E_1 \times E_2 \ldots \times E_n$, por ejemplo, si los conjuntos $E_1, \ldots, E_n$ están provistos de una adición, se podrá proveer E de la adición

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n).$$

Si e_i es el elemento neutro de E_i $(1 \leq i \leq n)$, la ley producto tendrá por elemento neutro $e = (e_1, e_2, \ldots, e_n)$; si $a_1, a_2, \ldots, a_n$ son simetrizables, respectivamente, en cada uno de los conjuntos E_i, $(a_1, a_2, \ldots, a_n)$ es simetrizable y

$$-(a_1, a_2, \ldots, a_n) = (-a_1, -a_2, \ldots, -a_n).$$

2. Relación de equivalencia compatible con una ley interna. Ley cociente

Consideremos un conjunto E sobre el cual están definidas:

— Una relación de equivalencia R: $x \equiv x'$ (mod R).

— Una ley de composición interna T.

Se puede estar tentado de definir una ley interna sobre E/R, es decir, entre dos clases $\dot{x}$, $\dot{y}$; la primera idea que se tiene es de considerar la clase [illegible]

$$(\dot{x}, \dot{y}) \to \widehat{x \top y}$$

sea una aplicación de E/R × E/R en E/R, es necesario que la clase $\widehat{x \top y}$ no dependa de los representantes x de $\dot{x}$ e y de $\dot{y}$, sino únicamente de las clases $\dot{x}$, $\dot{y}$, es decir, que

$$[x \equiv x' \text{ (mod R)} \quad \text{e} \quad y \equiv y' \text{ (mod R)}] \Rightarrow x \top y \equiv x' \top y' \text{ (mod R)}$$

En este caso se dice que *la relación de equivalencia* R *es compatible con la ley* T. Se podrá entonces definir una ley de composición T entre las clases poniendo

$$(\dot{x}) \top (\dot{y}) = \left(\dot{\widehat{x \top y}}\right),$$

puesto que la clase segundo miembro es independiente de los representantes escogidos x e y y no depende, pues, más que de las clases $\dot{x}$ e $\dot{y}$.

La ley interna T definida sobre E/R se llama *la ley cociente de la ley* T *por la relación de equivalencia* R. La aplicación de E en E/R definida por

$$x \to \dot{x}$$

es, por definición suprayectiva; se le llama *la suprayección canónica de* E sobre E/R, en general no es inyectiva (lo es si y sólo si toda clase $\dot{x}$ contiene un solo elemento, es decir, si R es la igualdad).

Si no da lugar a confusiones, se podrá designar la ley T y la ley cociente T por el mismo símbolo T.

Se ve fácilmente que si la ley T es asociativa o conmutativa, lo es también la ley cociente; si la ley T tiene un elemento neutro e, la ley cociente tiene por elemento neutro $\dot{e}$, si a tiene un simétrico a' para la ley T $\dot{a}$ es simetrizable y tiene por simétrico $\dot{a}'$ para la ley cociente; pero la ley cociente puede tener un elemento neutro sin que la ley T lo tenga y $\dot{a}$ puede ser simetrizable en E/R, sin que ningún representante de la clase $\dot{a}$ lo sea para la ley T; ahí radica uno de los intereses de una ley cociente. Contrariamente, a puede ser regular para T y $\dot{a}$ no serlo para la ley cociente (ver ej. 2 más abajo).

La noción de compatibilidad de una relación de equivalencia con una ley de composición interna puede de hecho descomponerse en dos:

Se dirá que la relación R *es compatible por la izquierda* (resp. *compatible por la derecha*) con la ley T definida sobre E si para todo y de E

$$x \equiv x' \pmod{R} \Rightarrow y \top x \equiv y \top x' \pmod{R}$$
$$(\text{resp. } x \equiv x' \pmod{R} \Rightarrow x \top y \equiv x' \top y \pmod{R}),$$

es decir, una relación R es compatible por la izquierda (resp. por la derecha) con una ley T si componiendo por la izquierda (resp. por la derecha) los dos miembros de una congruencia módulo R, se obtiene también una congruencia módulo R.

Después de lo que hemos visto anteriormente, *una relación de equivalencia* R *es compatible con una ley* T *si y sólo si es compatible por la izquierda y por la derecha con* T.

EJERCICIOS

1. Demostrar que la relación p divide $x - x'$ en **Z** (p entero natural > 1) es compatible con la adición y la multiplicación en **Z**. Construir las tablas de adición y de multiplicación de los conjuntos $\mathbf{Z}/p\mathbf{Z}$ para $p = 2$, $p = 3$, $p = 5$, $p = 6$ (ver § 18, ej. 3).

2. Con la ayuda de la tabla de multiplicar del ejercicio precedente para $p = 6$ demostrar que a puede ser regular para la ley definida sobre E y $\dot{a}$ no serlo para la ley cociente.

3. Demostrar que si p es un número primo, todo elemento $\dot{x} \neq 0$ de $\mathbf{Z}/p\mathbf{Z}$ es inversible (utilizar la fórmula de Bezout: $ux + vp = 1$; ver § 99, t. 5).

54. Ley interna definida sobre $\mathfrak{F}$(E, F) deducida de una ley interna definida sobre F

Siendo E un conjunto cualquiera y F un conjunto provisto de una ley interna T, consideremos dos aplicaciones f y g de E en F, es decir, dos elementos de $\mathfrak{F}$(E, F); podemos definir una ley interna $\perp$ sobre $\mathfrak{F}$(E, F) poniendo

$$h = f \perp g \Leftrightarrow (\forall x \in E) \qquad h(x) = f(x) \top g(x).$$

Se ve que si T es asociativa y conmutativa, también lo es ⊥. Si T posee un elemento neutro e, la función constante que toma este valor e para todo x de E es elemento neutro para la ley ⊥.

En general no hay ningún inconveniente en representar estas dos leyes sobre F y $\mathfrak{F}(E, F)$ por el mismo símbolo.

EJEMPLO

Sea $E = F = \mathbf{R}$, sobre $\mathfrak{F}(\mathbf{R}, \mathbf{R})$, se pondrá

$$s = f + g \Leftrightarrow (\forall x \in \mathbf{R}) \quad s(x) = f(x) + g(x)$$
$$p = fg \Leftrightarrow (\forall x \in \mathbf{R}) \quad p(x) = f(x)g(x).$$

s y p son, respectivamente, la *suma* y el *producto* de dos funciones f y g.

OBSERVACION

Si $E = F$ y si sobre E está definido un producto, no se confundirá las dos leyes definidas sobre $\mathfrak{F}(E, E)$, respectivamente, por

$$(f, g) \mapsto f \circ g, \qquad (f, g) \mapsto fg$$

(ver observación 2, § 15).

EJERCICIOS

1. Demostrar que para la ley $f + g$ definida sobre $\mathfrak{F}(\mathbf{R}, \mathbf{R})$ hay un elemento neutro que se determinará y que todo elemento f posee un opuesto que se determinará.
2. Demostrar que para la ley fg definida sobre $\mathfrak{F}(\mathbf{R}, \mathbf{R})$ hay un elemento neutro que se determinará. *¿Cuáles son los elementos f regulares, los elementos f simetrizables?*

12. Transporte de una ley interna por biyección

Sea E un conjunto provisto de una ley interna T y f una biyección de E sobre un conjunto E′. Podemos definir una ley interna T′ sobre E′ de la siguiente manera natural: cualesquiera que sean x' e y' de E′, existe x e y únicos de E, tales que

$$f(x) = x' \qquad f(y) = y'$$

por definición, el elemento compuesto de x' e y' por la ley T′ será $f(x \top y)$, es decir,

(1) $\qquad (\forall x', y' \in E') \qquad x' \top' y' = f[f^{-1}(x') \top f^{-1}(y')].$

Se dice que *la ley* T′ *de* E′ *ha sido definido por transporte de la ley* T *de* E *mediante la biyección* f *de* E *sobre* E′. La relación (1) puede también escribirse

(2) $\qquad (\forall x, y \in E) \qquad f(x \top y) = f(x) \top' f(y).$

El estudio de las relaciones tales como (2), sea f biyectiva o no, se hará en la sección III; referimos al lector a ella para el estudio de las propiedades de la ley T′ deducidas de las propiedades de la ley T.

III. Homomorfismos e isomorfismos de (E, T) en (E′, T′)

56. Homomorfismo de (E, T) en (E′, T′)

Consideremos una aplicación f de (E, T) en (E′, T′), se puede formar el compuesto de x e y en E, sea $x \top y$ y su imagen $f(x \top y)$ en E′; se puede formar también las imágenes $f(x)$ y $f(y)$, después su elemento compuesto en E′, sea $f(x) \top' f(y)$. Las aplicaciones de (E, T) en (E′, T′) tales que estos dos elementos de E′ sean iguales desempeñan un gran papel en Algebra, de donde la importancia de la siguiente definición:

DEFINICIÓN. — *Se llama* homomorfismo *de* (E, T) *en* (E′, T′) *toda aplicación f de* E *en* E′ *tal que*

$$(\forall x, y \in E) \qquad f(x \top y) = f(x) \top' f(y).$$

Un homomorfismo de (E, T) *en* (E, T) *se llama un* endomorfismo *de* (E, T).

EJEMPLOS

1. Siendo R una relación de equivalencia definida sobre E compatible con la ley interna definida sobre E, la aplicación canónica de E sobre E/R (ver § 53) es un homomorfismo de (E, T) sobre (E/R, †), pues para toda pareja x, y de E

$$\dot{\widehat{x \top y}} = \dot{x} \dagger \dot{y}$$

se le dice el *homomorfismo canónico* de (E, T) sobre (E/R, †).

2. La aplicación f de (N, +) en (N, ×) definida por

$$x \to f(x) = 2^x$$

es un homomorfismo, pues $2^{x+y} = 2^x 2^y$.

3. Igualmente (ver curso de Análisis) la aplicación de (R, +) en (R, ×) definida por $(a > 0)$

$$x \to f(x) = a^x$$

es un homomorfismo, pues $a^{x+y} = a^x a^y$.

4. La aplicación de $(\mathbf{R}^*_+, \times)$ en (R, +) definida por $(a > 0)$

$$x \mapsto g(x) = \log_a x$$

es un homomorfismo, pues $\log_a (xy) = \log_a x + \log_a y$.

Un homomorfismo de (E, T) en (E′, T′) se dice *suprayectivo*[10], *inyectivo*[10], *biyectivo* según que la aplicación f sea suprayectiva, inyectiva, biyectiva.

(10) Ciertos autores llaman epimorfismo a un homomorfismo suprayectivo y monomorfismo a un homomorfismo inyectivo.

EJERCICIO

Entre los homomorfismos dados como ejemplos, determinar los que son suprayectivos, inyectivos, biyectivos.

TEOREMA 1. — *Si f es un homomorfismo de* (E, T) *en* (E', T') *y g es un homomorfismo de* (E', T') *en* (E'', T''), $g \circ f$ *es un homomorfismo de* (E, T) *en* (E'', T'').

Cualesquiera que sean x e y de E tenemos

$$\begin{aligned}(g \circ f)(x \mathrm{T} y) &= g[f(x \mathrm{T} y)] && \text{(definición de } g \circ f)\\ &= g[f(x) \mathrm{T}' f(y)] && (f \text{ es un homomorfismo})\\ &= g[f(x)] \mathrm{T}'' g[f(y)] && (g \text{ es un homomorfismo})\\ &= (g \circ f)(x) \mathrm{T}'' (g \circ f)(y) && \text{(definición de } g \circ f).\end{aligned}$$

TEOREMA 2. — *Si f es un homomorfismo de* (E, T) *en* (E', T'), *entonces:*

a) f(E) *es una parte estable de* E' *para la ley* T'.

b) Si la ley T *es asociativa (resp. conmutativa), la ley inducida por la ley* T' *sobre* f(E) *es asociativa (resp. conmutativa).*

c) Si e es elemento neutro de (E, T), $e' = f(e)$ *es elemento neutro de* (f(E), T'). *Si, además, x es simetrizable en* (E, T) *y admite x_1 por simétrico, $f(x)$ es simetrizable y admite $f(x_1)$ por simétrico en* (f(E), T').

Este teorema enuncia propiedades de la ley T' sobre f(E) deducidas de propiedades de la ley T sobre E. Consideremos, pues, los *elementos arbitrarios* x', y', z' de f(E), existe x, y, z de E tales que

$$f(x) = x', \qquad f(y) = y', \qquad f(z) = z'.$$

a) x e y pertenecen a E, análogamente $x \mathrm{T} y$, luego

$$f(x \mathrm{T} y) = f(x) \mathrm{T}' f(y) = x' \mathrm{T}' y' \in f(\mathrm{E}).$$

b) La relación de asociatividad aplicada a x, y, z en E

$$(x \mathrm{T} y) \mathrm{T} z = x \mathrm{T} (y \mathrm{T} z)$$

da en f(E) por el homomorfismo f

$$\begin{aligned}f[x \mathrm{T} (y \mathrm{T} z)] &= f(x) \mathrm{T}' f(y \mathrm{T} z) = f(x) \mathrm{T}' [f(y)' \mathrm{T}' f(z)] = x' \mathrm{T}' (y' \mathrm{T}' z')\\ &= f[(x \mathrm{T} y) \mathrm{T} z] = f(x \mathrm{T} y) \mathrm{T}' f(z) = [f(x) \mathrm{T}' f(y)] \mathrm{T}' f(z) = (x' \mathrm{T}' y') \mathrm{T}' z'.\end{aligned}$$

Igualmente

$$x \mathrm{T} y = y \mathrm{T} x \Rightarrow f(x \mathrm{T} y) = f(y \mathrm{T} x) \Rightarrow x' \mathrm{T}' y' = y' \mathrm{T}' x'.$$

c) $e \mathrm{T} x = x \mathrm{T} e = x \Rightarrow f(e \mathrm{T} x) = f(x \mathrm{T} e) = f(x)$, es decir, poniendo

$$e' = f(e): \; e' \mathrm{T}' x' = x' \mathrm{T}' e' = x'.$$

Además, sea x_1, simétrico de x supuesto simetrizable,

$$x \mathrm{T} x_1 = x_1 \mathrm{T} x = e \Rightarrow f(x \mathrm{T} x_1) = f(x_1 \mathrm{T} x) = f(e),$$

es decir,

$$f(x) \mathrm{T}' f(x_1) = f(x_1) \mathrm{T}' f(x) = e'.$$

Vemos, pues, que todas estas propiedades de la ley T′ *han sido demostradas únicamente para la ley inducida sobre* $f(\mathrm{E})$; no se puede decir nada para la ley T′ sobre E′ — $f(\mathrm{E})$, puesto que la igualdad fundamental

$$f(x \mathrm{T} y) = f(x) \mathrm{T}' f(y)$$

sólo concierne los elementos de E y de $f(\mathrm{E})$.

Por otra parte, de las propiedades de la ley T′ (incluso sobre $f(\mathrm{E})$), no podremos en general deducir nada para la ley T; sea, en efecto, x', y', z', los elementos de $f(\mathrm{E})$, imágenes respectivas de x, y, z, de E, de una igualdad en $f(\mathrm{E})$ tal como $z' = x' \mathrm{T}' y'$, no podemos en general deducir que z y $x \mathrm{T} y$ son iguales, sino solamente que por el homomorfismo f, z y $x \mathrm{T} y$ tienen la misma imagen en $f(\mathrm{E})$. Hay, sin embargo, un caso en que se puede relevar las propiedades de (E′, T′) a las de (E, T), es el caso en que f es *biyectiva*. Consagraremos el párrafo siguiente al estudio de estos homomorfismos biyectivos.

57. Isomorfismo de (E, T) sobre (E′, T′)

Consideremos un homomorfismo biyectivo, f de (E, T) sobre (E′, T′), siendo f una biyección, cualesquiera que sean x' e y' de E′, existe x e y *únicos* de E imágenes recíprocas respectivas de x' e y'. Tenemos, pues, siendo f un *homomorfismo,*

$$f(x \mathrm{T} y) = f(x) \mathrm{T}' f(y) = x' \mathrm{T}' y'$$

y siendo f una biyección

$$f^{-1}(x' \mathrm{T}' y') = x \mathrm{T} y = f^{-1}(x') \mathrm{T} f^{-1}(y'),$$

luego f^{-1} es un homomorfismo (biyectivo) de (E′, T′) sobre (E, T), de donde:

TEOREMA 3 Y DEFINICIÓN. — *Si* f *es un homomorfismo biyectivo de* (E, T) *sobre* (E′, T′), f^{-1} *es un homomorfismo biyectivo de* (E′, T′) *sobre* (E, T); *se dice que* f *es un* isomorfismo *de* (E, T) *sobre* (E′, T′).

Un isomorfismo de (E, T) *sobre* (E, T) *se llama un* automorfismo *de* (E, T).

EJEMPLOS Y EJERCICIOS

1. Si se designa por D el conjunto descrito por 2^x ($x \in \mathbf{N}$), la aplicación f de (N, +) sobre (D, ×) definida por $x \to f(x) = 2^x$ es un isomorfismo.

2. f aplicación de $\mathbf{R}$ sobre $\mathbf{R}_+^*$ definida por ($a > 0$): $x \to a^x$ es un isomorfismo de ($\mathbf{R}$, +) sobre ($\mathbf{R}_+^*$, ×). ¿Cuál es el isomorfismo recíproco?

3. Sea E el conjunto de las rotaciones planas de centro O y de ángulo $k\frac{2\pi}{p}$ (k entero positivo, negativo o nulo cualquiera, p entero natural > 1 dado), mostrar que (E, O) es isomorfo a ($\mathbf{Z}/p\mathbf{Z}$, +).

4. Si la ley T′ está definida sobre E′ por transporte de la ley T definida sobre E mediante una biyección f de E sobre E′ (ver § 55), f es un isomorfismo de (E, T) sobre (E′, T′).

5. La aplicación $f = \mathrm{id}_E$ es un automorfismo de (E, T).

Naturalmente un isomorfismo posee todas las propiedades de un homomorfismo en particular:

COROLARIO DEL TEOREMA 1. — *El elemento compuesto $g \circ f$ de dos isomorfismos es un isomorfismo.*

Si consideramos la "relación" siguiente sobre la "*clase*" (ver §§ 3 y 20) de los conjuntos provistos de una ley interna: "*Existe un isomorfismo de* (E, T) *sobre* (E′, T′)", esta relación es:

- *Reflexiva,* pues id_E es un automorfismo de (E, T) (ver ej. 5).
- *Simétrica* (ver teorema 3).
- *Transitiva* (ver corolario anterior).

Esta "relación" que se expresa: "(E, T) *y* (E′, T′) *son isomorfos*" se designa por ciertos autores

$$(E, T) \approx (E', T')$$

o

$$E \approx E'$$

si no se teme ninguna ambigüedad sobre las leyes internas T y T′.

Pero el hecho de que el *homomorfismo f* sea *biyectivo* entraña resultados más precisos que los del teorema 2 (§ 56): De una manera general, toda propiedad de la ley T sobre E (resp. de la ley T′ sobre E′) implica la *misma propiedad* de la ley T′ sobre E′ (resp. de la ley T sobre E). Veremos múltiples ilustraciones de este hecho en lo sucesivo. Demos seguidamente un ejemplo:

EJERCICIO

6. Si f es un isomorfismo de (E, T) sobre (E′, T′), demostrar que si a es regular en (E, T), igualmente lo es $a' = f(a)$ en (E′, T′).

¿Valdría este resultado si f fuera un homomorfismo no biyectivo? (Considerar el homomorfismo canónico $\mathbf{Z} \mapsto \mathbf{Z}/n\mathbf{Z}$.)

IV. Simetrización de una ley interna - Grupo aditivo Z

58. Simetrización de la adición en N. Grupo aditivo Z

a) La adición en **N** es asociativa, conmutativa, posee un elemento neutro 0, todo elemento es regular, pero ningún elemento, salvo cero, es simetrizable; más generalmente la ecuación

$$x + a' = a \tag{1}$$

no tiene solución en N más que si $a \geq a'$, esta solución se escribe entonces $a - a'$. Busquemos un conjunto G provisto de una ley T conmutativa tal que N sea una parte estable de G para T, con la adición como la ley inducida sobre N y tal que todo elemento de G sea simetrizable. Convenimos designar igualmente por el signo + la ley en G. Entonces cualquiera que sea la pareja (a, a') de enteros naturales, la ecuación (1) será resoluble en G. Pero a un tal elemento $a - a'$ de G no corresponderá una pareja única (a, a'), sino todas las parejas (b, b') tales que

$$(2) \qquad b - b' = a - a'.$$

Pero esta relación es equivalente a

$$(2') \qquad a + b' = b + a' \quad \text{(relación R)}$$

que es una relación definida sobre N × N. Nos vemos así obligados a considerar el conjunto producto N × N provisto de la relación R.

b) Proveamos, además, al conjunto N × N de una ley producto (ver § 52), que expresaremos también por el signo +,

$$(a, a') + (b, b') = (a + b, a' + b').$$

Se ve fácilmente que sobre N × N, la adición es asociativa, conmutativa, todo elemento es regular y hay un elemento neutro (0, 0).

Consideremos ahora la relación R definida por (2'), es una equivalencia; en efecto:

— es *reflexiva*, pues para toda pareja (a, a')

$$a + a' = a + a',$$

— es *simétrica*, ya que

$$a + b' = b + a' \Rightarrow b + a' = a + b',$$

— es *transitiva*, pues

$$a + b' = b + a' \quad \text{y} \quad b + c' = c + b'$$

entrañan

$$(a + b') + c' = (b + a') + c' = (b + c') + a' = (c + b') + a',$$

de donde siendo regular todo elemento de N (simplificación por b'): $a + c' = c + a'$.

Designamos por $\left(\overset{\cdot}{\widehat{a, a'}}\right)$ la clase de (a, a') módulo R, tendremos para todo x de N

$$\left(\overset{\cdot}{\widehat{a, a'}}\right) = \left(\overset{\cdot}{\widehat{a + x, a' + x}}\right)$$

estas clases describen el conjunto $E = (N \times N)/R$, para que podamos definir una ley cociente sobre E, es necesario que la relación R sea compatible con la adición en $N \times N$ (ver § 53); ahora bien,

$$(a_1, a_1') \equiv (a_2, a_2') \quad \text{y} \quad (b_1, b_1) \equiv (b_2, b_2') \pmod{R}$$

entrañan

$$a_1 + a_2' = a_2 + a_1' \quad \text{y} \quad b_1 + b_2' = b_2 + b_1',$$

de donde por adición y teniendo en cuenta la asociatividad y la conmutatividad de la adición en N

$$(a_1 + a_2') + (b_1 + b_2') = (a_2 + a_1') + (b_2 + b_1')$$
$$(a_1 + b_1) + (a_2' + b_2') = (a_2 + b_2) + (a_1' + b_1),$$

luego

$$(a_1 + b_1, a_1' + b_1') \equiv (a_2 + b_2, a_2' + b_2') \pmod{R}$$

podemos, pues, poner por *definición* en E, anotando la ley cociente por el signo $\dot{+}$.

(3) $$\left(\overset{\cdot}{\widehat{a, a'}}\right) \dot{+} \left(\overset{\cdot}{\widehat{b, b'}}\right) = \left(\overset{\cdot}{\widehat{a + b, a' + b'}}\right).$$

La ley $\dot{+}$ así definida en E es asociativa y conmutativa, tiene un elemento neutro $\varepsilon = \left(\overset{\cdot}{\widehat{0, 0}}\right)$.

Como para todo x de N, $0 + x = x + 0$, este elemento neutro $\left(\overset{\cdot}{\widehat{0, 0}}\right)$ es igual a $\left(\overset{\cdot}{\widehat{x, x}}\right)$ y, además, todo elemento es simetrizable; en efecto,

$$\left(\overset{\cdot}{\widehat{a, a'}}\right) \dot{+} \left(\overset{\cdot}{\widehat{a', a}}\right) = \left(\overset{\cdot}{\widehat{a + a', a' + a}}\right) = \varepsilon,$$

luego

$$-\left(\overset{\cdot}{\widehat{a, a'}}\right) = \left(\overset{\cdot}{\widehat{a', a}}\right).$$

El conjunto E *está provisto de una ley interna (adición):*

— *Asociativa y conmutativa.*

— *Poseyendo un elemento neutro.*

— *Tal que todo elemento es simetrizable para esta ley.*

Diremos que (E, $\dot{+}$) *tiene una estructura de grupo conmutativo, o también que es un grupo conmutativo para la adición.*

c) Busquemos ahora cómo el grupo E nos permitirá resolver el problem planteado. Los elementos de E que corresponden a las soluciones en **N**

$x + a' = a$; es decir, las clases $\left(\overset{\cdot}{\widehat{a,\ a'}}\right)$ con $a \geq a'$ pueden escribirs $\left(\overset{\cdot}{\widehat{a - a',\ 0}}\right) = \left(\overset{\cdot}{\widehat{x,\ 0}}\right)$. Estos elementos describen una parte E′ de E q es estable para la adición, puesto que

$$(4) \qquad \left(\overset{\cdot}{\widehat{x,\ 0}}\right) \dot{+} \left(\overset{\cdot}{\widehat{y,\ 0}}\right) = \left(\overset{\cdot}{\widehat{x + y,\ 0}}\right).$$

Consideremos la aplicación f de (**N**, $+$) en (E′, $\dot{+}$) definida por

$$x \rightarrow f(x) = \left(\overset{\cdot}{\widehat{x,\ 0}}\right)$$

por definición de E′, f es *suprayectiva,* es también *inyectiva,* pu $\left(\overset{\cdot}{\widehat{x,\ 0}}\right) = \left(\overset{\cdot}{\widehat{y,\ 0}}\right)$ implica $x + 0 = 0 + y$, luego $x = y$, f es, pues, *biyectiv*

por otra parte, la igualdad (4) muestra que (**N**, $+$) y (E′, $\dot{+}$) son isomorfo lo que se puede esquematizar por la figura

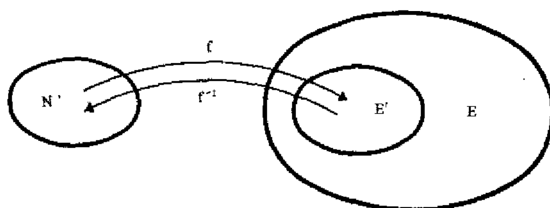


FIG. 9

No hemos resuelto completamente el problema propuesto al principio d párrafo; pero si escribimos las leyes en **N** y E (y E′) por el mismo símbolo + (**N**, $+$) y (E′, $+$) tienen las mismas propiedades (ver § 17), nos vemos con ducidos a *identificar* **N** y E′ poniendo *por convenio*

$$\left(\overset{\cdot}{\widehat{x,\ 0}}\right) = x,$$

los elementos $\left(\overset{\cdot}{\widehat{x,\ 0}}\right)$ no nos interesan por ellos mismos, sino únicament porque ellos verifican la relación (4) que tiene las mismas propiedades qu la adición en **N**. Gracias a esta *identificación,* **N** se encuentra dentro d grupo conmutativo E que se le designa por **Z**, los elementos de **Z** se les llam *enteros racionales.*

Pondremos $\mathbf{Z}^* = \mathbf{Z} - \{0\}$.

d) Representemos por las letras griegas los elementos de **Z**: la ecuación (α y α' elementos dados de **Z**)

$$\xi + \alpha' = \alpha$$

tiene siempre una solución $\xi = \alpha - \alpha'$. Si α y α' son enteros naturales a y a', tendremos

$$\xi + a' = a \leftrightarrow \xi + \left(\overline{a', 0}\right) = \left(\overline{a, 0}\right)$$

gracias a la identificación de más arriba, de donde

$$\xi = \left(\overline{a, 0}\right) + \left[-\left(\overline{a', 0}\right)\right] = \left(\overline{a, 0}\right) + \left(\overline{0, a'}\right) = \left(\overline{a, a'}\right)$$

elemento de **Z** que designaremos también $a - a'$.

Si $a \geq a'$, $a - a'$ pertenece a **N**, si $a' \geq a$, $a' - a$ pertenece a **N**, pues [illegible] es opuesto de $\left(\overline{a, a'}\right) = a - a'$, luego $a' - a$ pertenece a [illegible] descrita por los opuestos de los elementos de **N**); por tanto,

$$\mathbf{Z} = \mathbf{N} \cup (-\mathbf{N})$$

[illegible] que

$$\mathbf{N} \cap (-\mathbf{N}) = \{0\}.$$

La primera de estas dos fórmulas muestra que **Z** es un grupo *minimal* [illegible] conteniendo **N**, puesto que un tal grupo debe contener al menos todos los opuestos de los elementos de **N**, o sea, $-\mathbf{N}$.

Queda una última pregunta, ¿este grupo minimal es único? Al haber identificado los dos conjuntos isomorfos (**N**, +) y (E', +) podemos decir que todo grupo (G, ⊤) isomorfo a (**Z**, +) respondería a la pregunta. En el próximo artículo consideraremos la misma cuestión de una manera más general, y veremos que de hecho todo grupo *minimal* respondiendo a la pregunta está determinado, salvo un isomorfismo.

[illegible] Simetrización de una ley asociativa, conmutativa definida sobre E para la cual todo elemento es regular

a) Primer enunciado del problema de simetrización

Diremos que (G, ⊥) es un *grupo* si

— la ley ⊥ es *asociativa,*

— hay en (G, ⊥) un *elemento neutro,*

— *todo* elemento de (G, ⊥) es *simetrizable.*

Dado (E, T) buscamos simetrizar la ley T, es decir, nos planteamos el *problema* siguiente:

Se puede sumergir (E, T) *en* (G, ⊥) *de manera que:*

1. *Sea* (G, ⊥) *un grupo.*

2. *Sea* E *una parte estable de* G, *siendo la ley* T *definida sobre* E *la ley inducida por* ⊥ *sobre* E.

3. *Sea* G *minimal, es decir, ningún grupo* $G' \subset G$ *con* $G' \neq G$ *responde a la pregunta.*

b) Condiciones verificadas por (E, T)

Siendo E una parte estable de G, la ley inducida T sobre E debe ser *asociativa* (ver § 51); por otra parte, todo elemento de G que al ser simetrizable es regular (ver § 48, *b*), es evidente, en consecuencia, que *todo elemento* de E debe de ser *regular* para la ley T (inducida sobre E por la de G). Estas condiciones no son suficientes, pero vamos a ver que *si exigimos, además, a* G *ser un grupo conmutativo, el problema propuesto tiene solución* y su solución es única en un sentido que precisaremos. Resulta de ello que suponemos también que la ley inducida sobre E (la ley T) es conmutativa. Para simplificar la escritura notaremos las dos leyes T sobre E y ⊥ sobre G *aditivamente*.

c) Construcción de G

G deberá contener al menos todos los elementos de E, un elemento neutro (si E no contenía) y todos los simétricos de los elementos de E.

Por otra parte, siendo a y a' elementos de E, si $-a'$ designa el simétrico de un elemento a' de E (considerado como elemento de G), G debe contener los elementos de la forma $a + (-a') = a - a'$ y todas las sumas de los elementos de esta forma; observemos que, gracias a la conmutatividad y asociatividad, la suma de dos elementos $a - a'$ y $b - b'$ es de la misma forma

$$(a - a') + (b - b') = (a + b) - (a' + b').$$

Finalmente G debe contener al menos todos los elementos de la forma $a - a'$ (a y a' elementos de E), donde la sustracción es relativa a la adición a definir en G, luego la *determinación* de G comprende:

1. La determinación de los *elementos* de G.

2. La determinación de la *ley* para la cual G tiene una estructura de *grupo conmutativo.*

Como hemos visto en el párrafo precedente, a un tal elemento $a - a'$ de G no corresponde una única pareja de elementos de E × E, sino todas las parejas tales que

$$b - b' = a - a' \Leftrightarrow a + b' = b + a' \qquad \text{(relación R)}$$

constataríamos, como en el párrafo precedente, que R es una relación de equivalencia definida sobre E × E, que nos permite definir el conjunto cociente (E × E)/R descrito por las clases $\left(\overset{\cdot}{\widehat{a, a'}}\right)$.

Luego una parte de G *está descrita por elementos* $a - a'$, *en correspondencia biunívoca con los elementos* $\left(\overset{\cdot}{\widehat{a, a'}}\right)$ *de* (E × E)/R.

Desgraciadamente E no es una parte de (E × E)/R: existe solamente una inyección entre los elementos de E y ciertos elementos de (E × E)/R; a saber, las clases $\left(\overset{\cdot}{\widehat{a, a'}}\right)$ de las parejas (a, a') tales que la ecuación $a' + x = a$ tenga solución en E. Nos vemos así conducidos a modificar el enunciado del problema propuesto:

d) Enunciado modificado del problema de la simetrización

Dado un conjunto provisto de una adición asociativa, conmutativa para la cual todo elemento de E *es regular, determinar un grupo conmutativo* G *del que una parte estable* G′ *sea isomorfa a* E, *al tomar* G′ *provisto de la ley inducida por la de* G *y* E *provisto de la ley dada, además deberá ser* G *minimal.*

Resulta de este enunciado que si G es una solución del problema, todo grupo G_1 isomorfo a G será también una solución, pues la parte G_1' correspondiente a G′ por este isomorfismo será isomorfa a E (la composición de dos isomorfismos es también un isomorfismo, ver § 57), *luego en el mejor de los casos* G *quedará definido, salvo un isomorfismo.*

Según las consideraciones desarrolladas más arriba, G debe contener (E × E)/R (o un conjunto isomorfo). Pero si demostramos que (E × E)/R responde a las condiciones impuestas a G en el enunciado modificado, habremos demostrado que el problema así planteado tiene una solución única minimal (salvo un isomorfismo), puesto que el menor, por inclusión, de los conjuntos conteniendo (E × E)/R es el mismo (E × E)/R.

Basta repetir las consideraciones del párrafo precedente: la relación R es compatible con la adición en E × E, de donde

$$\left(\overset{\cdot}{\widehat{a, a'}}\right) + \left(\overset{\cdot}{\widehat{b, b'}}\right) = \left(\overset{\cdot}{\widehat{a + b,\ a' + b'}}\right).$$

Esta adición en (E × E)/R es asociativa, conmutativa, hay un elemento neutro $\varepsilon = \left(\overset{\cdot}{\widehat{x, x}}\right)$ (x elemento cualquiera de E, luego no es necesario que haya un elemento neutro en E). Todo elemento $\left(\overset{\cdot}{\widehat{a, a'}}\right)$ es simetrizable

$$-\left(\dot{\overline{a', a}}\right) = \left(\dot{\overline{a, a'}}\right),$$

luego $(E \times E)/R$ es un *grupo conmutativo.* Pongamos $G = (E \times E)/R$.

Demostremos que hay en G una parte estable G′ isomorfa a E, sabem
que está descrita por las clases $\left(\dot{\overline{a, a'}}\right)$ de las parejas (a, a') tales q
$a' + x = a$ tenga solución en E. Este elemento x determina la cl
$\left(\dot{\overline{x + z, z}}\right)$, siendo z un elemento cualquiera de E, pues para todo z y todo
de E

$$(x + z,\ z) \equiv (x + t,\ t) \qquad (\text{mod } R).$$

Recíprocamente una clase como $\left(\dot{\overline{x + z, z}}\right)$ de G′ determina un elemen
único x de E; en efecto,

$$\left(\dot{\overline{x + z, z}}\right) = \left(\dot{\overline{y + t, t}}\right) \Leftrightarrow (x + z) + t = z + (y + t),$$

es decir, $x = y$, al ser la ley de E asociativa, conmutativa y siendo to
elemento regular.

Luego la aplicación f de E en G′ definida por

$$x \rightarrow f(x) = \left(\dot{\overline{x + z, z}}\right)$$

es *biyectiva* (la complicación de la escritura en relación al párrafo pre
dente es debida a que no se ha supuesto la existencia de un elemento neu
0 en E); por otra parte,

$$\left(\dot{\overline{a + u, u}}\right) + \left(\dot{\overline{b + v, v}}\right) = \left(\dot{\overline{a + b + u + v, u + v}}\right),$$

es decir, $(G', +)$ es una *parte estable* de $(G, +)$ y cualesquiera que se
a y b de E

$$f(a + b) = f(a) + f(b),$$

luego f es un *isomorfismo* de $(E, +)$ sobre $(G', +)$.

Por consiguiente, $G = (E \times E)/R$ corresponde al enunciado modificado,
todo grupo que responde a la pregunta le debe ser isomorfo.

Como en el párrafo precedente, *identificando* G′ y E, es decir, ponien

$$\left(\dot{\overline{x + z, z}}\right) = x,$$

se puede decir que G responde a la pregunta propuesta en la primera form

TEOREMA Y DEFINICIÓN. — *Dado un conjunto* E *provisto de una ley* T *asociativa, conmutativa, para la cual todo elemento de* E *es regular, existe un grupo conmutativo* (G, T) *único, salvo un isomorfismo, tal que una parte estable* G′ *de* G *es isomorfa a* (E, T) *y cumpliendo la condición de ser* G *minimal.*
G *se llama simetrizado de* (E, T).
Identificando G′ *y* (E, T) *se dice que se ha* sumergido (E, T) *en el grupo* G.

Si no puede surgir ninguna confusión, se escribe algunas veces $G' = E$.

c) Ejemplo. Conjunto Q_+^* de los racionales estrictamente positivos

Aplicando las consideraciones precedentes a $(N^*, \times)$ que responde a las condiciones impuestas a (E, T): multiplicación asociativa, conmutativa, para la cual todo elemento es regular, el grupo G será $(N^* \times N^*)/R$, la relación R está definida por

$$(a, a') \equiv (b, b') \pmod{R} \Leftrightarrow ab' = ba'.$$

El elemento neutro de G es $\left(\dot{\overline{1,\,1}}\right) = \left(\dot{\overline{x,\,x}}\right)$ cualquiera que sea x de N^*

la inversa de $\left(\dot{\overline{a,\,a'}}\right)$ es $\left(\dot{\overline{a',\,a}}\right)$, pues

$$\left(\dot{\overline{a',\,a}}\right)\left(\dot{\overline{a,\,a'}}\right) = \left(\dot{\overline{aa',\,a'a}}\right) = \left(\dot{\overline{1,\,1}}\right).$$

La parte estable $(G', *)$ de $(G, *)$ isomorfa a $(N^*, \times)$ será descrita por las clases $\left(\dot{\overline{a,\,a'}}\right)$ de las parejas (a, a') tales que $a'x = a$ tenga solución en N^*, es decir, que a' divida a. $(N^*, *)$ posee el elemento neutro 1, estas clases pueden escribirse $\left(\dot{\overline{x,\,1}}\right)$.

Identificando $\left(\dot{\overline{x,\,1}}\right)$ de G′ y x de N^* se ve que la ecuación $a'\xi = a$ es soluble en G; en efecto, esta ecuación se escribe en G

$$\left(\dot{\overline{a',\,1}}\right)\xi = \left(\dot{\overline{a,\,1}}\right),$$

de donde

$$\xi = \left(\dot{\overline{a,\,1}}\right)\left[\left(\dot{\overline{a',\,1}}\right)\right]^{-1} = \left(\dot{\overline{a,\,1}}\right)\left(\dot{\overline{1,\,a'}}\right) = \left(\dot{\overline{a,\,a'}}\right).$$

Por un abuso de notación (ver § 26, *c*, ejemplo 2), se representa esta clase por $aa'^{-1} = a/a'$, *pero estrictamente* a/a' *es una fracción representante de la* $\left(\dot{\overline{a,\,a'}}\right)$ *que se llama un número racional estrictamente positivo.* G está representado por Q_+^*: conjunto de los racionales estrictamente positivos.

60. Grupo aditivo totalmente ordenado Z

a) La relación de orden total definido sobre **N** por $a \leq b$ es equivalente a $b - a \in \mathbf{N}$ y es compatible con la adición (ver § 33, *b*). Vamos a demostrar que se puede definir sobre **Z** un orden total único que anotaremos también $a \leq b$ tal que:

— Este orden sobre **Z** induce el orden ya definido sobre **N**.

— Que sea compatible con la adición en **Z**, es decir, tal que

$$(\forall c \in \mathbf{Z}) \qquad a \leq b \Rightarrow a + c \leq b + c.$$

En efecto, en **Z** tendremos

$$a \leq b \Rightarrow a + (-a) \leq b + (-a) \Leftrightarrow b - a \geq 0.$$

Consideremos *a priori* esta relación definida sobre **Z**

$$(1) \qquad b - a \in \mathbf{N}$$

induce la relación $a \leq b$ sobre **N**. Por otra parte,

— Esta relación es *reflexiva*, pues $a - a = 0 \in \mathbf{N}$.

— Es *antisimétrica*, pues $b - a$ y $a - b$ son opuestos y $\mathbf{N} \cap (-\mathbf{N}) = \{0\}$

$$(b - a \in \mathbf{N} \quad \text{y} \quad a - b \in \mathbf{N}) \Rightarrow b - a = 0.$$

— Es *transitiva*, pues

$$[b - a \in \mathbf{N} \quad \text{y} \quad c - b \in \mathbf{N}] \Rightarrow (b - a) + (c - b) = c - a \in \mathbf{N}.$$

La relación (1) es, por lo tanto, una *relación de orden* sobre **Z**; este orden es *total*, pues $\mathbf{Z} = \mathbf{N} \cup (-\mathbf{N})$ entraña que $b - a$ pertenezcan a **N** o a $-\mathbf{N}$, luego que $b - a$ o $a - b$ pertenezca a **N**.

En fin, para todo c de **Z** es

$$b - a \in \mathbf{N} \Rightarrow b - a = b + c - (a + c) \in \mathbf{N},$$

luego *la relación de orden* (1) *es compatible con la adición en* **Z**.

El orden definido sobre **Z** *por la relación* (1) *satisface, pues, a las condiciones puestas y es el único. Diremos que la adición y la relación de orden* $a \leq b$ *compatible con la adición confieren a* **Z** *una estructura de grupo aditivo totalmente ordenado.*

Los elementos de este grupo aditivo totalmente ordenado **Z** se llaman *enteros racionales*[11], los elementos de **N** *enteros positivos*[12], los de $\mathbf{N}^*$ *enteros estrictamente positivos*, los de $-\mathbf{N}$ *enteros negativos*[12] y los de $-\mathbf{N}^*$ *enteros estrictamente negativos*.

---

(11) Algunos autores les dicen enteros relativos.

(12) Se observará que 0 es a la vez positivo y negativo. Si se quiere evitar toda confusión con otras terminologías (donde positivo y negativo significa, respectivamente, elemento de $\mathbf{N}^*$ y elemento de $-\mathbf{N}^*$) se podrá designar a los elementos de **N** por «positivos en sentido amplio» y los de $(-\mathbf{N})$ por «negativos en sentido amplio» (ver § 21, *c*).

Por otra parte (§ 36), cualesquiera que sean a y b estrictamente positivos, existe n de **N** tal que $na > b$; diremos que **Z** es un grupo totalmente ordenado *arquimediano.*

b) Se llama *valor absoluto* la aplicación de **Z** en **N** definida por

$$a \to |a| = \sup(a, -a)$$

por *abuso de lenguaje* se dice que $|a|$ es el *valor absoluto* de a

$$\begin{aligned} a &\geq 0 \Leftrightarrow |a| = a \\ a &\leq 0 \Leftrightarrow |a| = -a \\ |a| &= 0 \Leftrightarrow a = 0. \end{aligned}$$

Se demostrará, como ejercicio, que cualesquiera que sean a y b de **Z**

$$\big||a| - |b|\big| \leq |a + b| \leq |a| + |b|$$

y cualesquiera que sean $a_1, a_2, \ldots, a_n$ de **Z**

$$|a_1 + a_2 + \ldots + a_n| \leq |a_1| + |a_2| + \ldots + |a_n|.$$

V. Conjunto provisto de dos leyes de composición interna. Distributividad. Anillo Z

41. Distributividad

DEFINICIÓN. — *Dadas dos leyes internas* $\top$ *y* $\perp$ *definidas sobre* E, *se dice que la ley* $\top$ *es* distributiva por la izquierda (*resp.* por la derecha) *en relación a la ley* $\perp$ *si*

$$\begin{aligned} &(\forall\ a, b, c \in \mathrm{E}) & a \top (b \perp c) &= (a \top b) \perp (a \top c) \\ &[\text{resp. } (\forall\ (a, b, c) \in \mathrm{E}) & (b \perp c) \top a &= (b \top a) \perp (c \top a)]. \end{aligned}$$

Se dice que la ley $\top$ *es* distributiva *con relación a la ley* $\perp$ *si es a la vez distributiva por la izquierda y distributiva por la derecha.*

En particular, si la ley $\top$ es conmutativa las tres nociones de distributividad están confundidas. Si la ley $\top$ está escrita multiplicativamente y la ley $\perp$ aditivamente, las relaciones de antes se escribirán

$$\begin{aligned} &(\forall\, a, b, c \in \mathrm{E}) & a(b + c) &= ab + ac \\ &(\forall\, a, b, c \in \mathrm{E}) & (b + c)a &= ba + ca. \end{aligned}$$

EJEMPLOS Y EJERCICIOS

1. En **N**, **Z**, **Q**, **R**, **C** la multiplicación es distributiva con relación a la suma.
2. En $\mathfrak{P}$(E) cada una de las leyes $\cup$ e $\cap$ es distributiva con relación a la otra.
3. En **Z** demostrar que la adición es distributiva con relación a las dos leyes inf (a, b) y sup (a, b). ¿Sucede igual para la multiplicación? Estudiar el caso en que se considera **N** en lugar de **Z**.

62. Anillo totalmente ordenado Z

a) Tratemos de definir en **Z** una multiplicación:

— Distributiva con relación a la adición.

— Y que prolongue la que está definida sobre **N** (ver § 34).

Designamos en este subpárrafo *a*) por $a, b \dots$ los elementos de **N** y p $-a, -b, \dots$ los elementos de $-$**N**; habiendo sido definido ab en **N**, n basta definir $(-a)b$, $b(-a)$ y $(-a)(-b)$, puesto que todo elemento de pertenece a **N** o a $-$**N**.

De la relación $a + (-a) = 0$ se deduce multiplicando por la derecha o p la izquierda por b y utilizando la distributividad

$$(1) \qquad [a + (-a)]b = 0\,b = 0 \Rightarrow ab + (-a)b = 0$$

$$(2) \qquad b[a + (-a)] = b\,0 = 0 \Rightarrow ba + b(-a) = 0,$$

de donde, por definición de un opuesto en **Z**,

$$(3) \qquad (-a)b = b(-a) = -(ab)$$

esta igualdad nos permite escribir

$$(4) \quad [a + (-a)](-b) = 0(-b) = -(0b) = 0 \Rightarrow a(-b) + (-a)(-b) = 0$$

de donde

$$(5) \qquad (-a)(-b) = -[a(-b)] = -(-ab) = ab.$$

Las condiciones dadas determinan, pues, una multiplicación en **Z**. Las pr piedades de la multiplicación en **N** y las igualdades (3) y (5) demuestran q esta *multiplicación en* **Z** *es asociativa, conmutativa, y admite* 1 *por elemen neutro.* Demostremos en fin que esta multiplicación es también *distributi* con relación a la adición (pues en la definición de la multiplicación con ayuda de (1), (2) y (4) no hemos utilizado esta distributividad más que los casos particulares). Examinando todos los casos posibles y utilizando definición de la multiplicación en **Z** y el hecho que en **N** la multiplicaci es distributiva con relación a la adición y a la sustracción (ver § 34), co probamos que en **Z** la multiplicación es distributiva con relación a la adició sea, por ejemplo,

$$(-a)(b - c) \qquad \text{con} \qquad b < c$$

$$(-a)(b - c) = -[(-a)(c - b)] = a(c - b) = ac - ab = (-a)b + (-a)(-c)$$

Además, el producto de dos elementos de **Z** es positivo si y sólo si l dos factores son los dos positivos o los dos negativos.

b) Luego **Z** está provisto de dos leyes internas, la adición y la multiplicación, con las propiedades

$A_1(\forall a, b, c \in \mathbf{Z})$	$(a+b)+c=a+(b+c)$	
$A_2(\exists 0 \in \mathbf{Z})(\forall a \in \mathbf{Z})$	$a+0=0+a=a$	
$A_3(\forall a \in \mathbf{Z})(\exists a' \in \mathbf{Z})$	$a+a'=a'+a=0$	$(a'=-a)$
$A_4(\forall a, b \in \mathbf{Z})$	$a+b=b+a$	
$A_5(\forall a, b, c \in \mathbf{Z})$	$(ab)c=a(bc)$	
$A_6(\exists 1 \in \mathbf{Z})(\forall a \in \mathbf{Z})$	$a1=1a=a$	
$A_7(\forall a, b \in \mathbf{Z})$	$ab=ba$	
$A_8(\forall a, b, c \in \mathbf{Z})$	$a(b+c)=ab+ac$	

Resumiremos estas propiedades diciendo que **Z** para la adición y la multiplicación posee una *estructura de anillo conmutativo provisto de un elemento unidad.*

Como en **N** se definirá la suma y el producto de una sucesión finita de elementos de **Z**; vemos que el producto $a_1, a_2, \ldots, a_n$ es positivo si y sólo si contiene un número par de factores negativos. Por otra parte, se demostrará por recurrencia que

$$|a_1 a_2 \ldots a_n| = |a_1|\ |a_2| \ldots |a_n|$$

y la equivalencia de $|a|=0$ y $a=0$ entraña que: *Un producto en* **Z** *es nulo si y sólo si al menos uno de los factores es nulo.*

Se demostrará, en fin, como en **N**, que la multiplicación es distributiva en relación a la sustracción. De la distributividad con relación a la adición y a la sustracción resultan las reglas de cálculo conocidas bajo el nombre de *"reglas de los paréntesis"*.

c) Hemos visto que la relación de orden total $a \leq b$ es compatible con la adición en **Z**. Para la multiplicación tenemos

$$(a \geq 0 \quad \text{y} \quad b \geq 0) \Rightarrow ab \geq 0,$$

de donde resulta que

$$(a \leq b \quad \text{y} \quad c \geq 0) \Rightarrow (b-a)c \geq 0,$$

es decir, $ac \leq bc$; dicho de otra manera: *en* **Z** *la relación de orden total* $a \leq b$ *es compatible con la adición y con la multiplicación por un entero positivo.*

Se expresa este hecho diciendo que la adición, la multiplicación y la relación de orden total $a \leq b$ definidas sobre **Z** le confieren una *estructura de anillo totalmente ordenado.*

En fin, el orden total sobre **Z**, siendo arquimediano (§ 60), diremos que **Z** es anillo totalmente ordenado *arquimediano.*

d) Si a es un entero racional cualquiera y m es un entero natural no nulo, definiremos a^m mediante

$$a^1 = a \qquad a^m = a^{m-1}a$$

se ve que $0^m = 0$; por otra parte:

Si m es impar: a y a^m son simultáneamente positivos o simultáneamente negativos.

Si m es par: a^m es siempre positivo.

Se demostrará fácilmente que para todo entero racional a, todo entero racional b y para todo entero natural m y todo entero natural n, se tiene

$$a^m a^n = a^{m+n}, \qquad (ab)^m = a^m b^m, \qquad (a^m)^n = a^{mn}.$$

63. Relación de divisibilidad en el anillo Z

DEFINICIÓN. — *Dados dos enteros racionales a y b, si existe un entero racional q tal que $a = bq$, se dice indistintamente:*

— b divide a o b es un divisor de a.

— a es divisible por b o a es un múltiplo entero de b.

Observemos que q es único si y sólo si $b \neq 0$, si $b = 0$ la relación precedente entraña $a = 0$ y q arbitrario.

Esta relación "b divide a" se escribe $b \mid a$.

Las relaciones

$$a = 1a = (-1)(-a),$$

muestran que a es siempre divisible por a, $-a$, 1 y -1.

Busquemos en qué condición $b \mid a$ y $a \mid b$; existe entonces q y q' tales que $a = bq$ y $b = aq'$, de donde $a = aqq'$; luego si $a \neq 0$: $qq' = 1$; nos vemos así conducidos a buscar los elementos inversibles de **Z**. La relación

$$|qq'| = |q| \qquad |q'| = 1$$

muestra que:

los únicos elementos inversibles de **Z** *son* 1 *y* -1 *y si a y b no son nulos*

$$(a \mid b \quad \text{y} \quad b \mid a) \Rightarrow b = \pm a.$$

Las consideraciones precedentes muestran que la relación $b \mid a$ *no es una relación de orden*: no es antisimétrica, se ve fácilmente que es reflexiva y transitiva.

EJERCICIO

En **Z** la relación $a \mid b$ es una relación de preorden (ver ej. 23, fin del capítulo 1). ¿Cuál es la relación de orden asociada?

64. División euclídea en Z. Aplicaciones

a) Dados dos enteros racionales a y b ($b > 0$), busquemos si se puede definir un entero racional único q tal que

$$(1) \qquad qb \leq a < (q+1)b.$$

Si q existe, es *único*; en efecto, es el mayor elemento de la parte de **Z** descrita por el entero racional m tal que $mb \leq a$.

Hemos demostrado la existencia de q si $a \geq 0$ (ver § 36); supongamos $a < 0$, $-a$ es positivo; existe, pues, q' positivo tal que

$$q'b \leq -a < (q' + 1)b$$

si $q'b = -a$ ponemos $q = -q'$, se tiene $qb = a$,
si $q'b < -a$ ponemos $q = -(q' + 1)$, se tiene entonces

$$-(q + 1)b < -a < -qb,$$

de donde, en todos los casos,

$$qb \leq a < (q + 1)b.$$

Resulta que si se pone $r = a - bq$

$$0 \leq r < b$$

y r es evidentemente único, de donde:

TEOREMA Y DEFINICIÓN. — *Dada una pareja de enteros racionales a y b ($b > 0$), existe una pareja única de enteros racionales q y r tales que*

$$a = bq + r \qquad \text{y} \qquad 0 \leq r < b.$$

Esta aplicación de $\mathbf{Z} \times \mathbf{N}^*$ *en sí mismo se llama* división euclídea, *q es el* cociente *y r el* resto *de esta división euclídea.*

Si $r = 0$, a es divisible por b, q es entonces el cociente (simplemente) de a por b.

b) De la definición de la división euclídea de a por b ($b > 0$) se deduce, siendo c estrictamente positivo y d un divisor común estrictamente positivo de a y b,

$$ac = (bc)q + rc \qquad 0 \leq rc < bc$$

y

$$a' = b'q + r' \qquad 0 \leq r' < b'$$

escribiendo $a = a'd$, $b = b'd$, $r = a - bq = (a' - b'q)d = r'd$.

La teoría de la divisibilidad en **Z** sigue de la existencia de la división euclídea; la expondremos en la sección III del capítulo 5, después de haber definido y estudiado la noción del ideal.

c) Consideremos la relación definida sobre **Z**, siendo p un entero estrictamente positivo: $p \mid x - y$ (esta relación ha sido ya objeto de ejemplos y de ejercicios en los §§ 18 y 53).

Es una relación de equivalencia: en efecto, puede escribirse "existe k de $\mathbf{Z}$ tal que $x - y = kp$", ella es *reflexiva*, pues $x - x = 0p$, *simétrica*, pues

$$x - y = kp \Rightarrow y - x = (-k)p,$$

transitiva, pues

$$[x - y = kp \quad \text{e} \quad y - z = k'p] \Rightarrow x - z = (k + k')p.$$

Sea x un entero racional cualquiera, existe una pareja única q, r de enteros racionales tales que

$$x = pq + r \qquad 0 \leq r < p,$$

luego

$$\dot{x} = \dot{r}$$

y r puede tomar p valores $0, 1, \ldots, p - 1$; hay, pues, en el conjunto cociente representado $\mathbf{Z}/p\mathbf{Z}$, p clases

$$\mathbf{Z}/p\mathbf{Z} = \left\{ \dot{0}, \dot{1}, \dot{2}, \ldots, \dot{\widehat{p-1}} \right\}$$

(explicaremos esta notación del conjunto cociente en el § 75).

La relación $p \mid x - y$ en general se expresa por

$$x \equiv y \pmod{p},$$

que se enuncia "*x congruente con y, módulo p*" y $\dot{x}$ se llama *un entero módulo p*.

Se ve que esta relación de equivalencia es compatible con la adición y la multiplicación en $\mathbf{Z}$; en efecto (h y k enteros racionales),

$$\left\{ \begin{array}{l} x' = x + hp \\ \quad\text{e} \\ y' = y + kp \end{array} \right. \Rightarrow \left\{ \begin{array}{l} x' + y' = x + y + (h + k)p \\ \\ x'y' = xy + (hy + kx + hk)p \end{array} \right.$$

se puede, pues, poner por definición (ver § 53)

$$\dot{x} + \dot{y} = \dot{\widehat{x + y}}, \qquad \dot{x}\dot{y} = \dot{\widehat{xy}}$$

se verificará fácilmente que las propiedades A_1 a A_8 (enunciadas en el § 62, *b*) se verifican, luego $\mathbf{Z}/p\mathbf{Z}$ provisto de la adición y de la multiplicación que acabamos de definir tiene una estructura de anillo conmutativo provisto de un elemento unidad, $\dot{0}$ es el elemento cero y $\dot{1}$ el elemento unidad. Este anillo se llama *anillo de los enteros módulo p*.

EJERCICIOS

1. Si f es una aplicación polinomia (ver § 186, *d*) de coeficientes en $\mathbf{Z}$ de $\mathbf{Z}^n$ en $\mathbf{Z}$ definida por

$$(a_1, a_2, \ldots, a_n) \rightarrow f(a_1, a_2, \ldots, a_n)$$

y si, con b un entero estrictamente positivo, r_i es el resto de a_i en la división euclídea de a_i por b $(1 \leq i \leq n)$

$$f(a_1, a_2, \ldots, a_n) \equiv f(r_1, r_2, \ldots, r_n) \pmod{b}.$$

2. ¿Existe en $\mathbf{Z}/p\mathbf{Z}$ clases $\dot{x}$, $\dot{y}$ distintas del elemento $\dot{0}$ tales que $\dot{x}\dot{y} = \dot{0}$?

3. **Demostrar que dados a y b de $\mathbf{Z}$ $(b \neq 0)$, existe al menos una pareja (q, r) de enteros racionales tales que**

$$a = bq + r \qquad |r| < |b|.$$

¿Esta pareja es única?

VI. Leyes externas[13]

63. Leyes externas

DEFINICIÓN. — *Dado un conjunto* E *de elementos* a, b ... *y un conjunto* Ω *de elementos* α, β ... *se llama* ley de composición externa *entre elementos de* E *y elementos de* Ω, *toda aplicación* f *de* $\Omega \times$ E *en* E

$$(\alpha, a) \rightarrow f(\alpha, a)$$

$f(\alpha, a)$, *elemento de* E, *es el* elemento compuesto *de* α *y de* a *por la ley externa considerada.* Ω *se llama el* conjunto de los operadores *para la ley externa considerada.*

Se escribe en general $f(\alpha, a)$ por un signo $\alpha \cdot a$ o más simplemente αa, en este caso se dirá que la ley externa es una *multiplicación externa.* En ciertos casos se distinguirá $(\alpha, a) \rightarrow \alpha a$ (*multiplicación por la izquierda*) y $(\alpha, a) \rightarrow a\alpha$ (*multiplicación por la derecha*).

EJEMPLOS

1. Sea $\Omega = \mathbf{R}$ y E el conjunto de los vectores libres de la geometría elemental, la multiplicación de un vector libre $\vec{X}$ por un real α

$$(\alpha, \vec{X}) \rightarrow \alpha\vec{X}$$

es una ley externa.

2. Sea E un conjunto provisto de una ley interna representada aditivamente, multiplicativamente o por el signo $\top$ y $\Omega = \mathbf{N}^*$ (ver § 45), la aplicación representada, respectivamente,

$$(n, a) \rightarrow na \qquad \text{con} \qquad 1 \,.\, a = a$$
$$(n, a) \rightarrow a^n \qquad \text{con} \qquad a^1 = a$$
$$(n, a) \rightarrow \overset{n}{\top} a \qquad \text{con} \qquad \overset{1}{\top} a = a$$

es una ley externa.

(13) Los resultados de esta sección —*salvo* éstos del § 66— no se utilizarán hasta el capítulo 7.

3. Siendo E un conjunto provisto de una ley interna $\top$, las dos aplicaciones

$$(a, x) \to a \top x \qquad (a, x) \to x \top a$$

son dos leyes externas cuyo conjunto de operadores es el mismo E: *toda ley interna es, pues, un caso particular de una ley externa.*

Si x describe una parte F de E, se designará αF el conjunto descrito por αx (α elemento de Ω).

Siendo F una parte de E, se dice que es *estable* para la ley externa si para todo α de Ω: $\alpha F \subset F$.

Los conjuntos Ω y E pueden estar provistos de leyes internas, de donde se siguen posibles *relaciones* entre la ley externa y las distintas leyes internas. Designemos las que tendremos que utilizar:

Si el conjunto Ω está provisto de una ley interna asociativa $\top$, se dice que *la ley externa es asociativa con relación a la ley* $\top$ *de* Ω *si*

$$(\forall\, \alpha,\ \beta \in \Omega)\,(\forall\, a \in E) \qquad (\alpha \top \beta)a = \alpha(\beta a).$$

Si el conjunto E está provisto de una ley interna $\top$, se dice que *la ley externa es distributiva con relación a la ley interna* $\top$ *de* E *si*

$$(\forall\, \alpha \in \Omega)\,(\forall\, a,\ b \in E) \qquad \alpha(a \top b) = (\alpha a) \top (\alpha b).$$

Si el conjunto E está provisto de una ley interna $\top$ y Ω provisto de una ley interna $\perp$, se dice que *la ley externa es distributiva con relación a las dos leyes internas* $\top$ *y* $\perp$ *si*

$$(\forall\, \alpha,\ \beta \in \Omega)\,(\forall\, a \in E) \qquad (\alpha \perp \beta)a = (\alpha a) \top (\beta a).$$

EJEMPLOS Y EJERCICIOS

4. Para el ejemplo 1 de más arriba se verifica que

$$\alpha(\beta \overleftrightarrow{X}) = (\alpha\beta)\overrightarrow{X}, \qquad \alpha(\overrightarrow{X} + \overrightarrow{Y}) = \alpha\overrightarrow{X} + \alpha\overrightarrow{Y}, \qquad (\alpha + \beta)\overrightarrow{X} = \alpha\overrightarrow{X} + \beta\overrightarrow{X}.$$

5. Para el ejemplo 2 anterior la ley externa es asociativa con la multiplicación en $\mathbf{N}^*$

$$p(qa) = (pq)a, \qquad (a^p)^q = a^{pq}, \qquad \overset{p}{\top}\left(\overset{q}{\top} a\right) = \overset{pq}{\top} a$$

la ley externa es distributiva con relación a las dos leyes internas: la ley interna de E y la adición en $\mathbf{N}^*$

$$(p + q)a = pa + qa \qquad a^{p+q} = a^p a^q \qquad \overset{p+q}{\top} a = \left(\overset{p}{\top} a\right) \top \left(\overset{q}{\top} a\right).$$

¿En qué caso la ley externa es distributiva en relación a la adición en $\mathbf{N}^*$?

6. Siendo E el conjunto de las aplicaciones f de $\mathbf{R}^n$ en $\mathbf{R}$ indefinidamente derivable (ver curso de Análisis), Ω siendo el conjunto de los operadores derivación $\dfrac{\partial}{\partial x_i}$ $(1 \leq i \leq n)$ estudiar las propiedades de la ley externa

$$\left(\frac{\partial}{\partial x_i}, f\right) \to \frac{\partial f}{\partial x_i}.$$

66. Estudio de un caso particular: $(n, a) \to \overset{n}{\top} a$

Consideremos de nuevo el ejemplo 2 último (§ 65) suponiendo que E está provisto de una ley interna asociativa con un elemento neutro e, los elementos a (y eventualmente b) siendo *simetrizables* (esto será en particular el caso si (E, T) es un grupo); podemos definir entonces una ley externa en la que el conjunto de los operadores es $\mathbf{Z}$ poniendo, respectivamente, según la notación adoptada para la ley interna de E,

$$na = \begin{cases} a + a + \dots + a & (n \text{ términos}) & \text{si } n > 0 \\ e & & \text{si } n = 0 \\ (-n)(-a) & & \text{s } n \end{cases}$$

$$a^n = \begin{cases} a\ a \dots a & (n \text{ factores}) & \text{si } n > 0 \\ e & & \text{si } n = 0 \\ (a^{-1})^{-n} & & \text{si } n < 0 \end{cases}$$

$$\overset{n}{\top} a = \begin{cases} a \top a \top \dots \top a & (n \text{ términos}) & \text{si } n > 0 \\ e & & \text{si } n = 0 \\ \overset{n}{\top} a' & (a' \text{ simétrico de } a) & \text{si } n < 0. \end{cases}$$

Se verificará fácilmente que cualesquiera que sean a y b *simetrizables* de E y p y q de $\mathbf{Z}$, se tiene

$$p(qa) = (pq)a, \qquad (a^p)^q = a^{pq}, \qquad \overset{p}{\top}\left(\overset{q}{\top} a\right) = \overset{pq}{\top} a$$

$$(p + q)a = pa + qa, \qquad a^{p+q} = a^p a^q, \qquad \overset{p+q}{\top} a = \left(\overset{p}{\top} a\right) \top \left(\overset{q}{\top} a\right)$$

y suponiendo la adición conmutativa, para todo n de $\mathbf{Z}$ es

$$n(a + b) = na + nb;$$

por el contrario, las dos fórmulas siguientes no son exactas más que si a y b son *permutables*

$$(ab)^n = a^n b^n, \qquad \overset{n}{\top} (a \top b) = \left(\overset{n}{\top} a\right) \top \left(\overset{n}{\top} b\right)$$

esto se verificará en particular para todo par (a, b) si la ley interna E es conmutativa.

67. Homomorfismo, isomorfismo de dos conjuntos E, E′ provisto cada uno de una ley externa, con el mismo dominio de operadores

Sea E y E′ dos conjuntos provistos cada uno de una ley de composición externa (que representaremos multiplicativamente), siendo el dominio de operadores Ω el mismo para las dos leyes, diremos que una aplicación f de E en E′ es un homomorfismo para estas dos leyes si

$$(\forall x \in E)(\forall \alpha \in \Omega) \qquad f(\alpha x) = \alpha f(x).$$

Si E″ es un tercer conjunto provisto también de una ley externa (representada multiplicativamente), con Ω el dominio de los operadores también, si g es un homomorfismo de E′ en E″ para las leyes externas, se ve inmediatamente que $g \circ f$ es un homomorfismo de E en E″ para las leyes externas; en efecto,

$$(\forall x \in E)(\forall \alpha \in \Omega) \qquad (g \circ f)(\alpha x) = g[f(\alpha x)] = g[\alpha(f(x))]$$
$$= \alpha g[f(x)] = \alpha[(g \circ f)(x)];$$

por otra parte, $f(E)$ es una parte estable para la ley externa definida sobre E′: en efecto, sea x' un elemento de $f(E)$, existe x de E tal que $x' = f(x)$ para todo α de Ω tendremos, pues,

$$f(\alpha x) = \alpha f(x) = \alpha x' \in f(E).$$

En fin, si f es un *homomorfismo biyectivo* de E sobre E′ para las dos leyes externas, f^{-1} es un *homomorfismo* (biyectivo) de E′ sobre E, se dice que f es un *isomorfismo* de E sobre E′ para las dos leyes externas o también que E y E′ son *isomorfos* para estas dos leyes. En efecto, cualquiera que sea x' de E′, existe x única imagen recíproca de x'; tenemos, pues, para todo α de Ω

$$f(\alpha x) = \alpha f(x) = \alpha x'$$
$$\alpha x = f^{-1}(\alpha x') \Leftrightarrow f^{-1}(\alpha x') = \alpha f^{-1}(x').$$

VII. Estructuras. Isomorfismos. Homomorfismos[14]

68. Noción de estructura

a) Las propiedades de una ley interna definida sobre un conjunto E le dan una estructura, por ejemplo:

—Un conjunto G provisto de una ley interna tiene una estructura de *grupo* para esta ley si:

(14) Esta sección se podrá leer rápidamente en una primera lectura, si se desea; pero tendrá que meditarse después del estudio de los capítulos 4, 5, 7.

— la ley es *asociativa,*

— está provista de un *elemento neutro,*

— *todo* elemento es *simetrizable,*

si, además, la ley es conmutativa, se dice que G tiene una estructura de *grupo conmutativo o abeliano.*

Se pueden definir las estructuras mediante varias leyes internas y externas, por ejemplo:

— Un conjunto A provisto de dos leyes internas adición y multiplicación posee una *estructura de cuerpo* para estas dos leyes si:

— A posee una estructura de *grupo conmutativo para la adición,*

— la multiplicación es *asociativa,*

— la multiplicación es *distributiva* con relación a la adición,

si, además, la ley es conmutativa, A está provisto de una estructura de *anillo conmutativo.*

— Un conjunto K provisto de dos leyes internas adición y multiplicación posee una *estructura de cuerpo* para estas dos leyes si:

— K posee una estructura de *anillo para estas dos leyes,*

— $K^* = K - \{e\}$ (e elemento neutro de la adición) posee una estructura de *grupo para la multiplicación,*

si, además, la multiplicación es conmutativa, K posee una estructura de *cuerpo conmutativo.*

— Un conjunto E provisto de una adición interna y de una multiplicación externa, siendo el conjunto de los operadores un cuerpo conmutativo K, posee una *estructura de espacio vectorial sobre* K si:

— E posee una estructura de *grupo conmutativo* para la adición interna,

— la multiplicación externa verifica cualesquiera que sean a y b de E y α, β de K

$$\alpha(\beta a) = (\alpha\beta)a$$
$$\varepsilon a = a \qquad (\varepsilon \text{ elemento neutro de la multiplicación en K})$$
$$(\alpha + \beta)a = \alpha a + \beta a$$
$$\alpha(a + b) = \alpha a + \alpha b.$$

— Un conjunto E provisto de dos operaciones internas, suma y multiplicación y de una operación externa, siendo el conjunto de los operadores un cuerpo conmutativo K, posee una *estructura de álgebra sobre* K si:

— E posee una estructura de *anillo* para las dos operaciones internas,

— E posee una estructura de *espacio vectorial sobre* K para la adición interna y la operación externa,

— para todo α de K y para todo a y todo b de E

$$\alpha(ab) = (\alpha a)b = a(\alpha b).$$

b) Estas estructuras, que son las que estudiaremos a lo largo de este curso, se llaman *estructuras algebraicas*: ellas están definidas al dar sobre E una o varias leyes internas y externas (que se llaman también *operaciones algebraicas*). Hay otras estructuras, por ejemplo, la estructura de orden (ver § 20), la estructura de espacio métrico (ver § 110, *b*).

Una estructura está, pues, definida sobre un conjunto E al dar *relaciones* entre elementos de E o entre elementos de E y entre elementos de otros conjuntos. E se llama el *soporte* de la estructura considerada.

Se ve que un conjunto en sí mismo no posee ninguna estructura, mientras no se definan sobre el mismo operaciones algebraicas, relaciones de orden, etc. Por ejemplo, una estructura de grupo es un par (G, T), donde la ley T posee las propiedades indicadas más arriba. Por abuso de lenguaje, en lugar de decir "*G posee una estructura de grupo para la ley* T", se dirá "G *es un grupo para la ley* T", o también si no da lugar a confusión "G *es un grupo*", entendiéndose que el lector sabe que G es un grupo para una ley determinada: por ejemplo, se dirá el grupo **Z** (entendiéndose: con la adición como la ley interna).

c) Un conjunto E puede estar provisto de *varias estructuras,* supongamos E provisto de dos estructuras S_1 y S_2, se podrá distinguir:

— El conjunto E.

— El conjunto E provisto de la estructura S_1.

— El conjunto E provisto de la estructura S_2.

— El conjunto E provisto de dos estructuras S_1 y S_2.

En este último caso, habrá que buscar las relaciones entre S_1 y S_2 y ver si son "*compatibles*" en un sentido que hay que determinar. Así, el conjunto **Z** provisto de la adición tiene una *estructura de grupo abeliano* (S_1), y provisto de la relación $a \leq b$, tiene una *estructura de orden total* (S_2); nosotros hemos visto que cualesquiera que sean a, b, c

$$a \leq b \Rightarrow a + c \leq b + c;$$

diremos entonces que las estructuras de grupo abeliano y de orden total son *compatibles* y que proveen **Z** de una *estructura de grupo abeliano totalmente ordenado.*

Hay que distinguir, pues:

— El conjunto soporte **Z**.

— El grupo abeliano **Z**,

— El conjunto totalmente ordenado **Z**.

— El grupo abeliano totalmente ordenado **Z**.

Evidentemente sería más racional utilizar símbolos diferentes para representar estos diversos seres matemáticos, pero ello complicaría las notaciones (así, **Z** posee aún una estructura de anillo conmutativa, una estructura de anillo totalmente ordenado, una estructura de anillo provisto de una división euclídea, etc.)

Cuando definamos un conjunto provisto de *varias estructuras* por los axiomas de estas estructuras, distinguiremos cuidadosamente los *axiomas de cada una de ellas* y los *axiomas de compatibilidad* entre ellas; por ejemplo, para una estructura de anillo habrá axiomas relativos a la adición (estructura de grupo abeliano), el axioma relativo a la multiplicación (asociatividad), los axiomas de "compatibilidad" (distribución por la izquierda y por la derecha de la multiplicación en relación con la suma).

Señalemos, en fin, que cuando se construye un conjunto E, sucede a menudo que este conjunto esté definido poseyendo una estructura S; así, **Z** está definido como grupo conmutativo minimal para la adición, tal que **N** sea una parte estable (ver § 58). Resulta fácil imaginar el conjunto soporte E y proveerle en seguida de otras estructuras.

59. Homomorfismos

a) En las páginas precedentes, hemos visto que en lugar de ocuparnos únicamente de números, de puntos o de vectores del espacio, o bien de funciones reales de variable real, como en matemáticas elementales, *nos hemos ocupado principalmente de las relaciones entre los elementos de un conjunto* E (*o de varios conjuntos*), *es decir, de las estructuras definidas sobre* E. Si consideramos entonces una aplicación f de E provisto de una estructura algebraica S en E′ provisto de una estructura algebraica S′, es interesante saber si la aplicación f transforma las relaciones que definen la estructura S, en las relaciones que definen la estructura S′. Esta cuestión nos conducirá a la noción de *estructuras algebraicas homólogas* y a la de *homomorfismo* de estas dos estructuras.

b) *Dados dos conjuntos* E *y* E′ *provistos de leyes internas y externas, diremos que estos dos conjuntos están provistos de estructuras algebraicas homólogas, si:*

– *A cada ley interna* ⊤ *definida sobre* E *corresponde una y sólo una ley interna* ⊤′ *definida sobre* E′.

– *A cada ley externa* ⊥ *definida sobre* E *corresponde una y sólo una ley externa* ⊥′ *definida sobre* E′, *y el dominio de operadores para las dos leyes* ⊥ *y* ⊥′ *el mismo.*

Así, **N**, **Z**, $\mathbf{Z}/p\mathbf{Z}$ provistos de la multiplicación y de la suma tienen estructuras homólogas.

Si, además, las propiedades fundamentales (asociatividad, conmutatividad, existencia de un elemento neutro, existencia de un simétrico, distributividad) son las mismas para los pares de leyes correspondientes, se dirá que las estructuras S y S′ de las que están provistos, respectivamente, E y E′ son de la misma *especie*: así, las estructuras de **N** y de **Z** para la adición y la multiplicación no son de la misma especie; por el contrario, las estructuras de **Z** y $\mathbf{Z}/p\mathbf{Z}$ son de la misma especie: *la especie de estructura algebraica llamada estructura de anillo conmutativo unitario.* Naturalmente dos conjuntos provistos de dos estructuras de la misma especie podrán tener las propiedades distintas para conceptos distintos de los conceptos fundamentales enumerados más arriba: así, **Z** es infinito y $ab = 0$ implica $a = 0$ o $b = 0$, mientras que $\mathbf{Z}/p\mathbf{Z}$ es finito y $\dot{a}\dot{b} = \dot{0}$ puede ser verdadero con $\dot{a}$ y $\dot{b}$ distintos de 0 (ver ej. 1, 2, § 53).

c) Estando E *y* E′ *provistos de dos estructuras homólogas* S *y* S′, *diremos que una aplicación f de* E *provista de* S *en* E′ *provista de* S′ *es un homomorfismo si para cada par de leyes internas correspondientes*

$$(1) \qquad (\forall x,\ y \in E) \qquad f(x \top y) = f(x) \top' f(y)$$

y si para cada par de leyes externas correspondientes (siendo el mismo el dominio de operadores de estas dos leyes)

$$(2) \qquad (\forall x \in E) \qquad (\forall \alpha \in \Omega) \qquad f(\alpha x) = \alpha f(x).$$

Del estudio hecho en los §§ 56 y 57 (leyes internas) y en el § 67 (leyes externas), se siguen los resultados siguientes:

—*El elemento compuesto* $g \circ f$ *de dos homomorfismos es un homomorfismo.*

—$f(E)$ *es una parte estable de* E′ *para la estructura* S′.

—*Si f es un homomorfismo biyectivo,* f^{-1} *es también un homomorfismo biyectivo, se dice que es un isomorfismo de* E *provisto de la estructura* S, *sobre* E′ *provisto de la estructura* S′; *se dice también que* E *provisto de* S *y* E′ *provisto de* S′ *son isomorfos.*

En particular, si las leyes que definen la estructura S′ sobre E′ han sido obtenidas por *transporte* de las leyes que definen la estructura S sobre E mediante una biyección *f*, esta biyección es un isomorfismo de E (provisto de S) sobre E′ (provisto de S′) (ver § 55).

Las leyes correspondientes en un isomorfismo tienen exactamente las mismas propiedades, como ya hemos visto que se verifica en el caso de una sola ley. Interesándonos más las relaciones entre elementos que los mismos elementos, nos encontraremos en la necesidad de *identificar* dos conjuntos E provisto de S y E′ provisto de S′ cuando sean *isomorfos* (es lo que hemos hecho en el § 58, *c*).

Señalemos finalmente que un homomorfismo *f* de E (provisto de S) en sí mismo se llama *endomorfismo* de E (provisto de S), y *automorfismo* de E (provisto de S) si, además, *f* es biyectiva.

Ejercicios

42*. Si la ley $\top$ definida sobre E es asociativa:

a) Demostrar por recurrencia que $\underset{1 \leq k \leq p+q}{\top} a_k = \left(\underset{1 \leq i \leq p}{\top} a_i\right) \top \left(\underset{p+1 \leq j \leq p+q}{\top} a_j\right)$.

b) Siendo I un conjunto totalmente ordenado finito $\{I_1, \ldots, I_p\}$ una partición de I tal que $l < m$ y $\alpha \in I_l$, $\beta \in I_m$ implican $\alpha < \beta$, estando cada I_l ordenado por el orden inducido sobre I_l por el de I, demostrar por recurrencia

$$\underset{i \in I}{\top} a_i = \underset{1 \leq l \leq p}{\top}\left(\underset{\lambda \in I_l}{\top} a_\lambda\right).$$

43*. Siendo asociativa y conmutativa la ley $\top$ definida sobre E:

a) Demostrar por inducción que, siendo $i \to f(i)$ una permutación de $[1, n]$,

$$\underset{1 \leq i \leq n}{\top} a_i = \underset{1 \leq i \leq n}{\top} a_{f(i)}.$$

b) Si $\{I_1, I_2, \ldots, I_p\}$ es una partición del conjunto finito I, demostrar por recurrencia que

$$\underset{i \in I}{\top} a_i = \underset{1 \leq l \leq p}{\top}\left(\underset{\lambda \in I_l}{\top} a_\lambda\right).$$

44. Sobre **R** ya provisto de la multiplicación y de la suma, se define la ley $\top$

$$a \top b = ab + a + b.$$

a) Demostrar que $\top$ es asociativa y conmutativa, y que posee un elemento neutro. ¿Cuáles son los elementos simetrizables?

b) La ley $\top$ ¿es distributiva respecto a la multiplicación y a la adición?

45. Siendo k y k' dos números reales dados, se define sobre **R**, ya provisto de la adición y de la multiplicación, la familia de leyes internas $\top$: $a \top b = kab + k'(a + b)$; determinar entre estas leyes las que son asociativas.

46. Siendo E un conjunto cualquiera se define sobre $\mathfrak{P}(E \times E)$ la ley $\top$ siguiente: si Y y X son dos partes de $E \times E$, $Y \top X$ es la parte de $E \times E$ constituida por las parejas (a, b) de $E \times E$ tales que existe c de E verificando la propiedad $(a, c) \in X$ y $(c, b) \in Y$:

a) Demostrar que la ley $\top$ es asociativa.

b) Dada una relación binaria R definida sobre E de grafo G, la transitividad de R es equivalente a $G \top G \subset G$.

47. E está provisto de una ley asociativa representada multiplicativamente:

a) Se supone que *existe a* de E tal que la traslación por la izquierda γ_a sea suprayectiva y que *existe u* de E tal que $ua = a$. Demostrar que, para todo x de E, $ux = x$. (Considerar la solución y de $ay = x$).

b) Se supone que *existe a* de E tal que las traslaciones por la izquierda γ_a y por la derecha δ_a sean suprayectivas. Demostrar que existe un elemento neutro.

c) Se supone en fin que *para todo* a de E, γ_a y δ_a son suprayectivas; demostrar que todo elemento de E es inversible.

48*. Sea E *finito* provisto de una ley asociativa representada multiplicativamente poseyendo un elemento neutro e, demostrar que todo elemento regular es simetrizable. Con la ayuda de un contraejemplo, demostrar que este resultado es falso si E es infinito.

49. Sobre $E = \{a, b, c\}$ se define una ley interna representada multiplicativamente por las igualdades

$$a^2 = a, \qquad b^2 = b, \qquad c^2 = c, \qquad bc = cb = a, \qquad ca = ac = b, \qquad ab = ba = c.$$

a) Demostrar que esta ley conmutativa no es asociativa. Comparar $(ab)c$ y $a(cb)$.

b) Verificar que todas las traslaciones (por la izquierda y por la derecha) son biyectivas.

50. Sobre $E = \{e, a, b, c\}$ se define una ley interna representada multiplicativamente, para la que e es elemento neutro, estando, además, la ley definida por las igualdades

$$a^2 = b^2 = c^2 = e, \qquad bc = cb = a, \qquad ca = ac = b, \qquad ab = ba = c.$$

Demostrar que esta ley es conmutativa, asociativa y que todo elemento es inversible.

51. Un elemento de (E, $\top$) es *idempotente* si $a \top a = a$. Una ley es *idempotente* si todo x de E es idempotente.

a) Demostrar que si una ley $\top$ asociativa y conmutativa es idempotente, la relación $x \top y = y$ es una relación de orden $(x \prec y)$; demostrar que $x \top y = \sup\ (x,\ y)$.

b) En un retículo (V. capítulo 1, ej. 24) se pone

$$a \vee b = \sup (a, b) \qquad\qquad a \wedge b = \inf (a, b).$$

Demostrar que las leyes $\vee$ y $\wedge$ son asociativas, conmutativas e idempotentes. Estudiar los ejemplos del ejercicio 24.

52*. Sea E provisto de una ley representada multiplicativamente, se dice que e' es un *elemento neutro por la izquierda* (resp. e'' *elemento neutro por la derecha*) si para todo x de E: $e'x = x$ (resp. $xe'' = x$).
Dado x de E, si existe x' tal que $x'x = e'$ (resp. x'' tal que $xx'' = e''$), se dice que x' es *inverso por la izquierda para* e' (resp. x'' *es inverso por la derecha para* e''). Se supone que la ley definida sobre E es asociativa, tiene un elemento neutro por la izquierda e' y que todo elemento x posee un inverso por la izquierda x' para e'.

a) Demostrar que $xx' = e'$ (considerar el inverso por la izquierda de x').

b) Demostrar que e' es también un elemento neutro por la derecha.

c) Deducir que E posee un elemento neutro único y que todo elemento de E es inversible.

53. Se considera dos conjuntos E_1 y E_2 provisto cada uno de una ley interna (representada multiplicativamente) y de una relación de equivalencia R_i ($i = 1, 2$), compatible con la ley definida sobre E_i.
Demostrar que la relación $R = R_1 \times R_2$ (ver capítulo 1, ej. 17) es compatible con la ley producto definida sobre $E_1 \times E_2$.

54. Se considera un retículo (ver capítulo 1, ej. 24) provisto de dos leyes internas $\vee$ y $\wedge$ (ej. 51 más arriba); se dice que el retículo es distributivo cuando cada una de las leyes $\vee$ y $\wedge$ es distributiva respecto la otra.
Determinar entre los retículos considerados en el ejercicio 24 los que son distributivos.

55. En **R** provisto de la suma y de la multiplicación, se considera el conjunto E descrito por $x + y\sqrt{2}$, de donde x e y describen **Q**.

a) Demostrar que E es estable para la adición y la multiplicación. Demostrar que todo elemento no nulo de E tiene un inverso en E.

b) Se supone E provisto de la adición y de la multiplicación inducidas por las de **R**, demostrar que E es isomorfo a $\mathbf{Q} \times \mathbf{Q}$ provisto de la adición y de la multiplicación siguientes

$$(x, y) + (x', y') = (x + x', y + y')$$
$$(x, y)(x', y') \quad = (xx' + 2yy', xy' + yx').$$

c) En el estudio precedente se reemplaza $\sqrt{2}$ por $\sqrt{d}$ (d entero natural no nulo); ¿qué propiedad debe poseer d para que los resultados de los §§ **a)** y **b)** subsistan?

56. Se considera un conjunto E provisto de una ley interna $\top$ y el conjunto $\mathfrak{F}(E, E)$ de las aplicaciones de E en E provisto de las leyes $f \circ g$ y $f \top g$ (V. § 54).
Demostrar que la ley $f \circ g$ es distributiva por la derecha respecto a la ley $\top$, pero en general no lo es por la izquierda.
Estudiar el caso E = **R** provisto: **a)** de la adición; **b)** de la multiplicación.

57. Estando E provisto de dos leyes internas $\top_1$, $\top_2$ y E′ de dos leyes internas $\top_1'$, $\top_2'$ correspondiendo, respectivamente, a $\top_1$ y a $\top_2$, se considera un homomorfismo de (E, $\top_1$, $\top_2$) en (E′, $\top_1'$, $\top_2'$); demostrar que si $\top_1$ es distributiva por la izquierda (resp. por la derecha) respecto a $\top_2$ en E, ocurre lo mismo con las leyes inducidas, respectivamente por $\top_1'$ y $\top_2'$ en f(E).

58. Encontrar todos los endomorfismos y todos los automorfismos: **a)** del grupo **Z**; **b)** del anillo **Z**.

59. *Siendo d un entero positivo que no es cuadrado de otro entero, se designa por* $\mathbf{Z}[\sqrt{d}]$ el conjunto descrito por $x + y\sqrt{d}$ cuando x e y describen **Z**.

a) Demostrar que $\mathbf{Z}[\sqrt{d}]$ es una parte estable de **R** para la adición y la multiplicación.

b) Los conjuntos $\mathbf{Z}[\sqrt{2}]$ y $\mathbf{Z}[\sqrt{3}]$ están provistos de la adición y multiplicación inducidas sobre ellos por la adición y la multiplicación de **R**, demostrar que no existe ningún isomorfismo entre $\mathbf{Z}[\sqrt{2}]$ y $\mathbf{Z}[\sqrt{3}]$.

4

GRUPOS

I. Definición. Ejemplos. Primeras propiedades

70. Definición. Notaciones. Ejemplos

DEFINICIÓN. — *Un conjunto* G *provisto de una ley interna* T, *es decir, un par* (G, T) *es un* grupo *si:*

G$_1$) *la ley* T *es asociativa.*

G$_2$) *existe un elemento neutro en* (G, T).

G$_3$) *todo elemento de* (G, T) *es simetrizable.*

Se dice también que G posee una *estructura de grupo* para la ley T, o que G es un grupo para la ley T, o más simplemente que G es un *grupo* si no hay ninguna ambigüedad sobre la ley.

Si, además, la ley es conmutativa, se dice que el grupo es *conmutativo o abeliano.*

La ley T se llama la *ley del grupo,* el elemento neutro *e* de (G, T) el *elemento neutro del grupo.*

Si card G $= n$, se dice que G es un *grupo finito, de orden n.*

Se llama *elemento central c* de un grupo G todo elemento que permuta con todo elemento de grupo, el *centro* C de G es el conjunto de sus elementos centrales: no es nunca vacío, pues contiene al menos *e*.

Cuando la ley de grupo está representada *multiplicativamente* (resp. *aditivamente*), se dice algunas veces que el grupo es *multiplicativo* (resp. *aditivo*). Utilizaremos en general la notación multiplicativa, o la notación aditiva (recordemos que en este caso supondremos un grupo abeliano, ver § 46).

Las reglas de cálculo en un grupo se deducen inmediatamente de los resultados del § 66, las recordamos con los axiomas de un grupo en el cuadro siguiente en las diversas notaciones:

	Adición	Multiplicación	Ley $\top$
$(\forall a, b \in G)$ $(\forall a, b, c \in G)$ $(\exists e \in G)(\forall a \in G)$ $(\forall a \in G)(\exists a' \in G)$	$a + b \in G$ $(a + b) + c = a + (b + c)$ $a + e = e + a = a$ (e elemento cero) $a + a' = a' + a = e$ ($a' = -a$ opuesto de a)	$ab \in G$ $(ab)c = a(bc)$ $ae = ea = a$ (e elemento unidad) $aa' = a'a = e$ ($a' = a^{-1}$ inverso de a)	$a \top b \in G$ $(a \top b) \top c = a \top (b \top c)$ $a \top e = e \top a = a$ (e elemento neutro) $a \top a' = a' \top a = e$ (a' simétrico de a)
$(\forall a, b \in G)$	$a + b = b + a$	$ab = ba$ (si el grupo es *conmutativo* solamente)	$a \top b = b \top a$
$(\forall a, b \in G)$	$-(a + b) = (-a) + (-b)$	$(ab)^{-1} = b^{-1}a^{-1}$	$(a \top b)' = b' \top a'$
$(\forall n \in \mathbf{Z})(\forall a \in G)$ $n > 0$	$na = a + a + \ldots + a$ (n términos)	$a^n = aa \ldots a$ (n factores)	$\overset{n}{\top} a = a \top a \top \ldots \top a$ (n términos)
$n = 0$	$0a = e$	$a^0 = e$	$\overset{0}{\top} a = e$
$n < 0$	$na = (-n)(-a)$	$a^n = (a^{-1})^{(-n)}$	$\overset{n}{\top} a = \overset{-n}{\top} a'$
$(\forall m, n \in \mathbf{Z})(\forall a \in G)$	$(m + n)a = ma + na$ $m(na) = (mn)a$	$a^{m+n} = a^m a^n$ $(a^m)^n = a^{mn}$	$\overset{m+n}{\top} a = \left(\overset{m}{\top} a\right) \top \left(\overset{n}{\top} a\right)$ $\overset{m}{\top}\left(\overset{n}{\top} a\right) = \overset{mn}{\top} a$
$(\forall n \in \mathbf{Z})(\forall a, b \in G)$	$n(a + b) = na + nb$	$(ab)^n = a^n b^n$ (si el grupo es *conmutativo* solamente)	$\overset{n}{\top}(a \top b) = \left(\overset{n}{\top} a\right) \top \left(\overset{n}{\top} b\right)$

EJEMPLOS Y EJERCICIOS

Verificar que los conjuntos siguientes provistos de las operaciones indicadas son grupos:

1. **Z**, **Q**, **R**, **C** para la suma.
2. $\mathbf{Q}^*$, $\mathbf{R}^*$, $\mathbf{C}^*$, $\mathbf{Q}^*_+$, $\mathbf{R}^*_+$ para la multiplicación.
3. El conjunto de las traslaciones del plano (o del espacio) para la composición de las traslaciones.
4. El conjunto de las rotaciones de centro O y de ángulo cualquiera en el plano para la composición de las rotaciones.
5. El conjunto de las rotaciones de centro O y de ángulo $k2\pi/n$ (n entero natural > 0 fijado, k entero racional cualquiera) para la composición de las rotaciones.
6. El conjunto de las isometrías del plano (o del espacio) para la composición de las aplicaciones. El conjunto de isometrías positivas (desplazamientos) para la misma operación; ¿es cierto también para el conjunto de isometrías negativas? (antidesplazamientos).
7. El conjunto $\mathfrak{B}(\mathrm{E}, \mathrm{E})$ de las biyecciones de un conjunto E sobre sí mismo para la composición de las aplicaciones (grupo de las permutaciones de E).
8. $\mathbf{Z}/n\mathbf{Z}$ (§ 64, *c*) para la adición.
9. El conjunto de la identidad y de las tres simetrías respecto a los tres ejes de un triedro trirrectángulo para la composición de las aplicaciones.

71. Propiedad fundamental de un grupo

TEOREMA. — *Para todo elemento a de un grupo* G *las aplicaciones de* G *en* G *definidas por* (ver § 47)

$$x \to \gamma_a(x) = ax \qquad x \to \delta_a(x) = xa$$

son biyectivas.

En efecto, γ_a es inyectiva, pues teniendo a un inverso a^{-1}

$$ax = ay \Rightarrow a^{-1}(ax) = a^{-1}(ay) \Rightarrow (a^{-1}a)x = (a^{-1}a)y,$$

luego $x = y$; lo mismo para δ_a. Son suprayectivas, pues cualquiera que sea b de G

$$ax = b \Rightarrow a^{-1}(ax) = a^{-1}b \Rightarrow x = a^{-1}b$$
$$xa = b \Rightarrow (xa)a^{-1} = ba^{-1} \Rightarrow x = ba^{-1}.$$

Este teorema puede enunciarse en la forma equivalente siguiente:

Para todo par (a, b) *de elementos de un grupo* G, *las ecuaciones* $ax = b$ *y* $xa = b$ *tienen una solución única.*

En fin, γ_a y δ_a son inyectivas, se puede enunciar:

COROLARIO. — *En un grupo todo elemento es regular.*

Considerando la tabla de un grupo finito (§ 44), podemos decir, de una manera intuitiva, que cada elemento figura una vez y sólo una en cada línea y en cada columna; un cuadro poseyendo esta propiedad es un *cuadrado*

latino. Luego todo grupo finito tiene por tabla un cuadrado latino, pero la recíproca no es cierta (ver ej. 1 más abajo). Pero se demuestra que si la ley es asociativa y si, cualesquiera que sean a y b de E, $ax = b$ y $ay = b$ tienen al menos una solución en E, entonces es un grupo (ver ej. 66, fin del capítulo).

EJERCICIOS

	a	b	c
a	a	c	b
b	c	b	a
c	b	a	c

1. La ley definida en el cuadro adjunto, ¿es una ley de grupo? (se estudiará si posee un elemento neutro y si es asociativa).

2. Demostrar que, en la tabla de un grupo finito el elemento neutro está situado sobre la diagonal principal o bien ocupa posiciones simétricas respecto a esta última.

3. Demostrar que, en un grupo finito de orden par, hay un número impar de elementos de G iguales a su propio inverso y distintos de e.

4. Encontrar todos los grupos de 2, 3, 4 elementos (buscar primero los cuadrados latinos de 2, 3, 4 elementos $\{e, a\}$, $\{e, a, b\}$, $\{e, a, b, c\}$ y utilizar los ejercicios anteriores). En estos casos particulares se podrá dispensar de verificar la asociatividad observando que para $n = 3$ se encuentra un solo cuadrado latino idéntico a la tabla del grupo aditivo $\mathbf{Z}/3\mathbf{Z}$ y que para $n = 4$ se encuentran dos cuadrados latinos, uno idéntico a la tabla del grupo de las simetrías en relación a los tres ejes de un triedro trirrectángulo, el otro a la tabla del grupo aditivo $\mathbf{Z}/4\mathbf{Z}$ (ver § 70, ej. 8 y 9).

II. Subgrupos de un grupo

72. Definición. Subgrupos del grupo aditivo Z

a) DEFINICIÓN.—*Se llama* subgrupo H *de un grupo* G *toda parte estable no vacía de* G, *que es ella misma un grupo para la ley inducida sobre* H *por la ley de* G.

Todos los subgrupos de G tienen el mismo elemento neutro e que G; en efecto, sea e' el elemento neutro de H para la ley inducida, tendremos

$$\text{en G} \quad ee' = e'e = e'$$
$$\text{en H} \quad e'e' = e',$$

luego $ee' = e'e'$ y $e = e'$, puesto que en un grupo todo elemento es regular.

G y $\{e\}$ son subgrupos de G. Todo subgrupo de G, distinto de G y de $\{e\}$ se llama *subgrupo propio* de G.

b) Subgrupos del grupo aditivo Z

Si H no contiene más que 0, $H = \{0\}$ subgrupo trivial. Supongamos H no reducido al elemento cero, conteniendo $n \neq 0$, H contiene $-n \neq 0$, luego H contiene enteros estrictamente positivos; el conjunto de los enteros estrictamente positivos de H tiene un elemento mínimo, sea a (ver § 28); cualquiera que sea x de H, existe un par q y r de enteros tales que (ver § 64)

$$x = aq + r \qquad 0 \leq r < a$$

x y a perteneciendo al subgrupo H, asimismo pertenecen aq y $r = x - aq$, siendo a el mínimo entero de H estrictamente positivo, $r = 0$, luego $x = aq$, luego

los subgrupos de **Z** *son los conjuntos* $a\mathbf{Z}$.

c) Sea H una parte no vacía de G; para que H sea un subgrupo de G es necesario primero que H sea estable, es decir (ver § 51),

(1) $(x \in \mathrm{H}$ e $y \in \mathrm{H}) \Rightarrow xy \in \mathrm{H}$ o $\mathrm{HH} \subset \mathrm{H}$.

Es inútil verificar la asociatividad de la ley inducida, ella es consecuencia de la asociatividad de la ley de G.

Asimismo todo elemento x de H tiene un inverso x^{-1} en H, es decir,

(2) $x \in \mathrm{H} \Rightarrow x^{-1} \in \mathrm{H}$ o $\mathrm{H}^{-1} \subset \mathrm{H}$.

El hecho de pertenecer e a H resulta de la aplicación de (1) a x y x^{-1}.

Las dos condiciones (1) y (2) son necesarias y suficientes para que la parte no vacía H de G sea un subgrupo. Estas dos condiciones pueden ser reemplazadas por la *única* condición siguiente:

TEOREMA. — *Para que una parte no vacía* H *de un grupo* G *sea un subgrupo de* G, *es necesario y suficiente que*

(3) $(x \in \mathrm{H}$ *e* $y \in \mathrm{H}) \Rightarrow xy^{-1} \in \mathrm{H}$.

Poniendo $y = x$ en (3) se obtiene

$$xx^{-1} = e \in \mathrm{H};$$

poniendo $x = e$ en (3) se obtiene

$$(\forall\, y \in \mathrm{H}) \qquad y^{-1} \in \mathrm{H};$$

en fin, reemplazando y por y^{-1} en (3), se obtiene

$$x(y^{-1})^{-1} = xy \in \mathrm{H},$$

luego H es una parte estable conteniendo e y la inversa de cada uno de sus elementos es, pues, un subgrupo de G.

En representación aditiva la condición (3) se escribirá

(3′) $(x \in \mathrm{H}$ e $y \in \mathrm{H}) \Rightarrow x - y \in \mathrm{H}$.

EJEMPLOS Y EJERCICIOS

1. Cada uno de los grupos aditivos **Z**, **Q**, **R**, **C** es un subgrupo de todos los siguientes. Cada uno de los grupos multiplicativos **Q***, **R***, **C*** es un subgrupo de todos los siguientes.

2. El conjunto de las traslaciones del plano es un subgrupo del grupo de los desplazamientos del plano; ¿ocurre lo mismo con el conjunto de las rotaciones del plano?, ¿con el conjunto de las rotaciones de centro dado O?

3. El conjunto de los desplazamientos (isometrías positivas) es un subgrupo del grupo de las isometrías del plano (o del espacio); ¿ocurre lo mismo para el conjunto de isometrías negativas?

4. Encontrar los subgrupos de los grupos de orden 3 y 4 estudiados en el ejercicio 4 (ver § 71).

5. Encontrar todos los subgrupos de $\mathbf{Z}/6\mathbf{Z}$, de $\mathbf{Z}/n\mathbf{Z}$ (ver § 70, ej. 8).

6. Demostrar que el centro de un grupo G es un subgrupo de G.

7. El conjunto de las biyecciones de E sobre E que conservan un elemento determinado a de E es un subgrupo del grupo de las permutaciones de E (ver § 70, ej. 7). Análogamente para el conjunto de las biyecciones de E dejando una parte A de E invariante.

8. Demostrar que para todo subgrupo H de un grupo G: $H^2 = H$, $H^{-1} = H$, $HH^{-1} = H$.

71. Intersección de subgrupos. Subgrupo engendrado por una parte A de un grupo G

a) La intersección $H_1 \cap H_2$ de dos subgrupos de G no es nunca vacía, contiene siempre el elemento neutro e de G; por otro lado, si x e y pertenecen a H_1 y a H_2, xy^{-1} pertenece a H_1 y a H_2, luego a su intersección; de una manera más general, dada una familia $\mathcal{H}$ de subgrupos H de G, su intersección I (ver § 7)

$$I = \bigcap_{H \in \mathcal{H}} H$$

no es vacía, pues $e \in H$, y cualesquiera que sean x, y perteneciendo a H, xy^{-1} pertenece a H para todo H de $\mathcal{H}$, luego a I; luego

toda intersección de subgrupos de G *es un subgrupo de* G.

b) Consideremos en particular una parte no vacía A de G y la familia $\mathcal{A}$ de los subgrupos de G conteniendo A; esta familia no es vacía, pues G es un elemento de $\mathcal{A}$. La intersección I de esta familia de subgrupos $\mathcal{A}$ es un subgrupo de G y es evidentemente el menor (con la relación de orden correspondiente a la inclusión de conjuntos).

Este subgrupo I, *intersección de la familia de los subgrupos de* G *que contienen* A, *se llama el subgrupo de* G *engendrado por* A.

EJEMPLOS Y EJERCICIOS

1. ¿Cuál es la intersección de los subgrupos $2\mathbf{Z}$ y $3\mathbf{Z}$ del grupo aditivo $\mathbf{Z}$? La misma pregunta para $a\mathbf{Z}$ y $b\mathbf{Z}$ (a, b enteros racionales).

2. ¿Cuál es la intersección del conjunto de los desplazamientos del espacio que conservan un punto A y del conjunto de los desplazamientos del espacio conservando un punto $B \neq A$?

3. Si A es un subgrupo de G, ¿cuál es el subgrupo engendrado por A?

4. Si a es un elemento de un grupo G, ¿cuál es el subgrupo engendrado por $\{a\}$?

5. Demostrar que el subgrupo engendrado por A, parte no vacía de un grupo G, está descrito por los productos finitos $a_1, a_2, \ldots, a_n$, donde a_i es un elemento de A o el inverso de un elemento de A.

74. Clases según un subgrupo

Suponiendo definida una relación de equivalencia R sobre un grupo G busquemos las relaciones de equivalencia compatibles con la ley de grupo (v. § 53), con el fin de definir una ley interna sobre el conjunto cociente G/R.

a) Busquemos primero las relaciones de equivalencia R_g compatibles por la izquierda con la ley de grupo.

La relación

$$x \equiv y \pmod{R_g}$$

entraña (compatibilidad por la izquierda)

$$e = x^{-1}x \equiv x^{-1}y \pmod{R_g}$$

esta relación de equivalencia, *si existe*, es tal que

$$x \equiv y \pmod{R_g} \Leftrightarrow x^{-1}y \in \dot{e}.$$

Demostremos que la clase $\dot{e}$ de e, módulo R_g, es un subgrupo de G; en efecto,

$$a \equiv e \pmod{R_g} \Leftrightarrow e = a^{-1}a \equiv a^{-1} \pmod{R_g},$$

luego (transitividad de R_g)

$$\begin{cases} a \equiv e \pmod{R_g} \\ \quad\text{y} \\ b \equiv e \pmod{R_g} \end{cases} \Rightarrow \begin{cases} a^{-1} \equiv e \pmod{R_g} \\ \quad\text{y} \\ b^{-1} \equiv e \pmod{R_g} \end{cases} \Rightarrow a^{-1}b \equiv e \pmod{R_g}$$

luego, si existe una equivalencia compatible por la izquierda con la ley de grupo, es de la forma $x^{-1}y \in \dot{e}$ y $\dot{e}$ es un subgrupo de G.

Recíprocamente, sea H un subgrupo cualquiera de G; consideremos la relación $x^{-1}y \in \mathrm{H}$, es una relación de equivalencia; en efecto:

— Es *reflexiva*: para todo x de G, $x^{-1}x = e \in \mathrm{H}$.

— Es *simétrica,* pues

$$x^{-1}y \in \mathrm{H} \Rightarrow (x^{-1}y)^{-1} = y^{-1}(x^{-1})^{-1} = y^{-1}x \in \mathrm{H}.$$

— Es *transitiva,* pues

$$[x^{-1}y \in \mathrm{H} \quad \text{e} \quad y^{-1}z \in \mathrm{H}] \Rightarrow (x^{-1}y)(y^{-1}z) = x^{-1}z \in \mathrm{H}.$$

Por otra parte, esta relación es *compatible por la izquierda* con la ley de grupo; en efecto, cualesquiera que sean x, y, z de G

$$x^{-1}y \in \mathrm{H} \Rightarrow (zx)^{-1}(zy) = (x^{-1}z^{-1})(zy) = x^{-1}y \in \mathrm{H},$$

luego *toda relación de equivalencia* R_g *compatible por la izquierda con la ley de un grupo* G *es de la forma*

$$x^{-1}y \in \mathrm{H},$$

siendo H *un subgrupo cualquiera de* G.

Se demostraría igualmente que *toda relación de equivalencia* R_d *compatible por la derecha con la ley de un grupo* G *es de la forma*

$$yx^{-1} \in H,$$

siendo H *un subgrupo cualquiera de* G.

b) Determinemos las clases de módulo de estas relaciones: y pertenecerá a la clase de x módulo R_g si y sólo si existe z de H tal que

$$x^{-1}y = z \Leftrightarrow y = xz.$$

La clase de x módulo R_g será, pues, xH, igualmente la clase de x módulo R_d es Hx; estas clases xH y Hx se llaman, respectivamente, *las clases por la izquierda y las clases por la derecha módulo el subgrupo* H.

La clase de e (módulo R_g o R_d) es H.

La aplicación $z \to xz$ de H en xH es suprayectiva (según la definición de xH), es inyectiva (pues en un grupo todo elemento es regular), es, pues, una biyección; luego:

Todas las clases por la izquierda xH *(y las clases por la derecha* Hx*) según el subgrupo* H *tienen, pues, la misma potencia: la potencia de* H.

Supongamos en particular el grupo G *finito* y de orden n, el subgrupo H es, pues, finito (parte de un conjunto finito); sea p su orden, el conjunto de las clases por la izquierda (o de las clases por la derecha) es una partición de G; existe, pues, q tal que $n = pq$, q es el número de clases por la izquierda (o el número de clases por la derecha), luego:

TEOREMA.—*En un grupo finito el orden de un subgrupo cualquiera es un divisor del orden del grupo.*

75. Caso de un grupo conmutativo. Grupo cociente

Representemos la ley aditivamente, las dos relaciones de equivalencia R_g y R_d están confundidas; la relación R única así definida se expresará por

$$x - y \in H \qquad \text{o} \qquad x \equiv y \quad (\text{mod } H);$$

las clases de equivalencia $\dot{x}$ ... describen el conjunto G/R, que se representa G/H. Siendo esta relación R compatible con la ley de grupo, podemos definir una adición en G/H por la relación (ver § 53)

$$(1) \qquad \dot{x} \dot{+} \dot{y} = \widehat{x + y}^{\,\cdot}.$$

La aplicación f de (G, +) en (G/H, $\dot{+}$) definida por

$$x \to f(x) = \dot{x}$$

es, pues, en virtud de (1) un *homomorfismo*, por definición de G/H es *suprayectivo*, luego (ver § 56, teorema 2) f(G) = G/H está provisto de una ley $\dot{+}$ asociativa, hay un elemento neutro, todo elemento $\dot{x}$ tiene un simétrico $(-\dot{x})$;

dicho de otra manera, (G/H, $\dot{+}$) es un grupo, se le llama el grupo cocient[e]
de G por el subgrupo H.

En general se expresará igualmente la adición en G/H por la notación +.

EJEMPLOS

1. Consideremos, por ejemplo, el grupo abeliano **Z**; sus subgrupos son de la form[a]
$n\mathbf{Z}$; la relación $x - y \in n\mathbf{Z}$ se escribe

$$x \equiv y \pmod{n}.$$

Los grupos cocientes se representarán, pues, $\mathbf{Z}/n\mathbf{Z}$, tal es el origen de esta notació[n]
que hemos introducido o utilizado en los §§ 53 y 64.

Este grupo se llama *grupo aditivo de los enteros módulo n*.

2. Siendo a un real no nulo, $a\mathbf{Z}$ es un subgrupo del grupo aditivo **R**; la relació[n]
$x - y \in a\mathbf{Z}$ se expresa

$$x \equiv y \pmod{a},$$

el grupo cociente $\mathbf{R}/a\mathbf{Z}$ es llamado grupo aditivo de los reales módulo a.

Para $a = 2\pi$ se obtienen los reales módulo 2π utilizados en la medida de los ángulo[s]

76. Subgrupo distinguido. Grupo cociente por un subgrupo distinguido

Si G no es abeliano, las relaciones $x^{-1}y \in \mathrm{H}$ e $yx^{-1} \in \mathrm{H}$ son en general di[s]tintas y siendo cada una de estas relaciones o bien compatible por la izquierd[a] o bien compatible por la derecha con la ley del grupo, es imposible en gener[al] definir una ley cociente sobre G/R_g o G/R_d. Esto será posible, sin embarg[o] si H está escogido de manera que R_g y R_d sean equivalentes, es decir, si l[as] particiones que definen sobre G son las mismas, o también *si toda clase p[or] la izquierda es una clase por la derecha y recíprocamente*, es decir, si par[a] todo x de G

$$x\mathrm{H} = \mathrm{H}x \Leftrightarrow \mathrm{H} = x\mathrm{H}x^{-1} \Leftrightarrow \mathrm{H} = x^{-1}\mathrm{H}x.$$

Un tal subgrupo H se llama *subgrupo distinguido* (o también *subgrup[o] invariante*, veremos el origen de esta denominación en el § 77).

Todo grupo G posee subgrupos distinguidos; por ejemplo, $\{e\}$ y G.

En un grupo *abeliano* todo subgrupo es distinguido.

Consideremos un subgrupo distinguido H de G y la relación de equi[valencia] valencia R

$$x^{-1}y \in \mathrm{H} \Leftrightarrow yx^{-1} \in \mathrm{H},$$

se puede definir un conjunto cociente G/R que se representa G/H; siendo [R] compatible por la izquierda y por la derecha con la ley de grupo, podem[os] definir una operación en G/H (operación que representaremos también multi[-]plicativamente) por la igualdad (ver § 53)

$$\dot{x}\dot{y} = \widehat{xy}$$

como en el párrafo precedente (ver § 56, t. 2), se verá que la aplicació[n]

$$x \rightarrow f(x) = \dot{x}$$

de G *en* G/H es un *homomorfismo suprayectivo,* pues $G/H = f(G)$ es un grupo que se llama *grupo cociente* de G por el subgrupo distinguido H.

EJERCICIOS

1. Sea un grupo G y dos de sus subgrupos H y K tales que $K \subset H \subset G$. Demostrar que si K es subgrupo distinguido de H, no es en general subgrupo distinguido de G. (Se demuestra incluso que si K es subgrupo distinguido de H y H subgrupo distinguido de G, K no es en general subgrupo distinguido de G.)
2. Demostrar que el centro C de un grupo es un subgrupo distinguido.
3. Demostrar que la intersección de dos subgrupos distinguidos de G es un subgrupo distinguido de G.
4. Demostrar que en el grupo de los desplazamientos del plano, el subgrupo de las traslaciones es un subgrupo distinguido.
5. Si H es un subgrupo distinguido de orden p de un grupo G de orden n, ¿cuál es el orden de G/H?

III. Homomorfismos, isomorfismos de los grupos

77. Resultados generales

Las definiciones y teoremas de los §§ 56 y 57 nos permiten enunciar las definiciones y teoremas siguientes:

a) DEFINICIÓN. — *Una aplicación f de un grupo* G *en un grupo* G′ *es un* homomorfismo de grupos *si y sólo si*

$$(\forall x,\ y \in G) \qquad f(xy) = f(x)f(y).$$

Si f es biyectivo, se dice que f es un isomorfismo *de grupos. Un homomorfismo de* G *en sí mismo es un* endomorfismo *del grupo* G *y un isomorfismo de* G *sobre sí mismo es un* automorfismo *del grupo* G.

Por ejemplo, siendo H un subgrupo distinguido de un grupo G (o un subgrupo cualquiera de un grupo abeliano), la aplicación (ver §§ 75 y 76) de G sobre G/H definida por $x \to f(x) = \dot{x}$ es un *homomorfismo suprayectivo,* se le llama el *homomorfismo canónico* de G sobre G/H.

Por otra parte, siendo a un elemento fijo de un grupo G, la aplicación f_a de G en G definida por

$$x \to x' = f_a(x) = axa^{-1}$$

es un *automorfismo* de G. En efecto, para todo x' de G, multiplicando por la izquierda por a^{-1} y por la derecha por a

$$x' = axa^{-1} \Rightarrow x = a^{-1}x'a,$$

luego f_a es biyectivo. Cualesquiera que sean x e y de G

$$f_a(xy) = a(xy)a^{-1} = (axa^{-1})(aya^{-1}) = f_a(x)f_a(y),$$

luego f_a es un homomorfismo biyectivo de G sobre G; es, pues, un automorfismo: los automorfismos f_a se llaman *automorfismos interiores* de G (se

reducen a la identidad para los grupos abelianos).

La definición de un subgrupo distinguido H de G, $H = aHa^{-1}$ para todo a de G (ver § 76), muestra que:

Un subgrupo H *de* G *es distinguido si y sólo si* H *es invariante para todos los automorfismos interiores de* G.

Este resultado explica el nombre de *subgrupo invariante* que se da también a los subgrupos distinguidos. Se puede incluso demostrar un resultado más general (ver ej. 5).

b) TEOREMA 1. — *La composición de dos homomorfismos (resp. isomorfismos) de grupos es un homomorfismo (resp. isomorfismo) de grupos.*

TEOREMA 2. — *Si f es un homomorfismo del grupo* G *de elemento neutro* e *en el grupo* G′ *de elemento neutro* e' *entonces:*

1. $e' = f(e)$, $f(x^{-1}) = [f(x)]^{-1}$.
2. $f(G)$ *es un subgrupo de* G′ *llamado* imagen *de f y representado* Im f.
3. $N = f^{-1}(e')$ *es un subgrupo distinguido de* G, *se le llama el* núcleo *del homomorfismo y se representa por* Ker f.

1. Para todo x de G, $f(xe) = f(ex) = f(x)f(e) = f(e)f(x)$, el elemento único z' de G′ tal que $f(x)z' = z'f(x) = f(x)$, siendo e' se tiene $e' = f(e)$. La demostración de $f(x^{-1}) = [f(x)]^{-1}$ ha sido ya realizada en el § 56.

2. Sea x' e y' dos elementos cualesquiera de $f(G)$ y x e y las imágenes recíprocas respectivas de x' e y', x, y^{-1}, xy^{-1} son elementos de G, luego

$$f(xy^{-1}) = f(x)f(y^{-1}) = f(x)[f(y)]^{-1} = x'y'^{-1}$$

es un elemento de $f(G)$.

3. Sean x e y dos elementos cualesquiera de N,

$$f(xy^{-1}) = f(x)[f(y)]^{-1} = e'e'^{-1} = e',$$

luego N es un subgrupo de G.

Por otra parte, para todo x de N y a de G

$$f(axa^{-1}) = f(a)f(x)f(a^{-1}) = f(a)e'f(a^{-1}) = f(a)f(a^{-1}) = e',$$

luego N es un subgrupo distinguido de G.

Por otra parte, $f(x) = f(y)$ es equivalente a $f(xy^{-1}) = e'$, luego xy^{-1} pertenece al núcleo, de donde:

COROLARIO. — *El homomorfismo de grupos f:* $G \to G'$ *es inyectivo si y sólo si* $f^{-1}(e') = \{e\}$.

OBSERVACION

Es inútil suponer que G′ es un grupo, basta con que G′ esté provisto de una operación interna, ya que con ello el teorema 2 del § 56 permite afirmar que $f(G)$ es una parte estable de G′ y tiene una estructura de grupo para la ley inducida por la de G y f es un homomorfismo suprayectivo del grupo G sobre el grupo $f(G)$.

TEOREMA 3. — *Si f es un isomorfismo de un grupo* G *sobre un grupo* G′, f^{-1} *es un isomorfismo del grupo* G′ *sobre el grupo* G. *Se dice entonces que los grupos* G *y* G′ *son isomorfos.*

EJERCICIOS

1. Si a es un número real estrictamente positivo, demostrar que la aplicación $x \rightarrow a^x$ del grupo aditivo **R** en el grupo multiplicativo $\mathbf{R}^*_+$ es un isomorfismo.

2. Siendo n un entero estrictamente positivo, demostrar que el grupo de las rotaciones del plano de centro O y de ángulo $k\dfrac{2\pi}{n}$ (k entero racional cualquiera) es isomorfo al grupo aditivo $\mathbf{Z}/n\mathbf{Z}$.

3. **Z** es un subgrupo aditivo de **R**; estudiar el homomorfismo canónico de **R** sobre **R**/**Z**. Este grupo aditivo de los reales módulo 1 se llama *toro de una dimensión*, se le representa por **T**. Siendo a un número real no nulo, demostrar que el grupo $\mathbf{R}/a\mathbf{Z}$ (§ 75. ej. 2) es isomorfo a **T**.

4. Demostrar que el conjunto de los automorfismos interiores de un grupo G (ver más arriba, a) es un grupo G′ para la composición de aplicaciones. Demostrar que la aplicación φ de G en G′ definida por $\varphi(a) = f_a$ es un homomorfismo. Demostrar que φ es un isomorfismo si y sólo si el centro de G es $\{e\}$.

5. Demostrar que un subgrupo estable para todos los automorfismos interiores de G es distinguido.

6. Sea f: G $\mapsto$ G′ un homomorfismo de grupos. Demostrar que:

X es un subgrupo de G $\Rightarrow f(\mathrm{X})$ es un subgrupo de G′.
X es un subgrupo distinguido de G $\Rightarrow f(\mathrm{X})$ es un subgrupo distinguido de $f(\mathrm{G})$.
Y′ es un subgrupo de G′ $\Rightarrow f^{-1}(\mathrm{Y}')$ es un subgrupo de G.
Y′ es un subgrupo distinguido de $f(\mathrm{G}) \Rightarrow f^{-1}(\mathrm{Y}')$ es un subgrupo distinguido de G.

7. Sea G un grupo, demostrar que la relación entre x y x' de G «existe a de G tal que $x' = axa^{-1}$» es una relación de equivalencia definida sobre G; se dice que x y x' son *conjugados* en G. Demostrar que si H es un subgrupo de G también lo es de $\mathrm{H}' = a\mathrm{H}a^{-1}$; se dice que H y H′ son *subgrupos conjugados* en G.

78. Descomposición canónica de un homomorfismo de grupos

Sea f un homomorfismo de G en G′. Consideremos la descomposición canónica de f (ver § 19, b), debemos considerar la relación de equivalencia sobre G: $f(x) = f(y)$.

Pero esta relación es equivalente a

$$f(x)[f(y)]^{-1} = f(x)f(y^{-1}) = f(xy^{-1}) = e',$$

luego es equivalente a $xy^{-1} \in \mathrm{N}$ y, puesto que el núcleo de f es un subgrupo distinguido de G (teorema 2 del párrafo precedente), esta relación es compatible con la ley de G; tenemos, pues, la descomposición $f = i \circ b \circ s$

$$\mathrm{G} \xrightarrow{s} \mathrm{G/N} \xrightarrow{b} f(\mathrm{G}) \xrightarrow{i} \mathrm{G}'.$$

i, aplicación canónica de $f(\mathrm{G})$ en G′ definida por $x' \rightarrow x'$, es un *homomorfismo inyectivo* de los grupos $f(\mathrm{G})$ y G′.

Hemos visto que s definida por $s(x) = \dot{x}$ (§ 76) es un *homomorfismo suprayectivo* de G sobre G/N. Consideremos la biyección b, definida por

$$\dot{x} \to x' = b(\dot{x}) = f(x),$$

siendo x un representante cualquiera de la clase $\dot{x}$. Sea $\dot{y}$ otra clase e y uno de sus representantes, xy es un representante de $\dot{x}\dot{y} = \dot{\widehat{xy}}$, puesto que la relación es compatible con la ley del grupo, luego

$$b(\dot{x}\dot{y}) = b\left(\dot{\widehat{xy}}\right) = f(xy) = f(x)f(y) = b(\dot{x})b(\dot{y}),$$

luego b biyectiva es un homomorfismo; es, pues, un *isomorfismo* de G/N sobre f(G) y está ligado de manera canónica a f:

TEOREMA. — *Todo homomorfismo f de un grupo G en un grupo G′ se descompone en,*

a) *El homomorfismo canónico de G sobre G/N, siendo N el núcleo de f.*
b) *El isomorfismo canónico de G/N sobre f(G).*
c) *El homomorfismo canónico de f(G) en G′.*

IV. Producto cartesiano de grupos. Suma directa

79. Producto cartesiano de grupos

Dados dos grupos (G_1, T_1), (G_2, T_2) podemos definir una ley interna T sobre el conjunto producto $G_1 \times G_2$ de una manera natural para la igualdad siguiente (ver § 52)

(1) $$(x_1, x_2)\,T\,(y_1, y_2) = (x_1\,T_1\,y_1,\ x_2\,T_2\,y_2)$$

si no hay lugar a confusión, las tres leyes podrán estar representadas de la misma manera; por ejemplo, en representación aditiva, o multiplicativa

(2) $$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1,\ x_2 + y_2)$$

(3) $$(x_1, x_2)(y_1, y_2) = (x_1y_1,\ x_2y_2)$$

a partir de ahora utilizaremos esta última notación.

Se ve fácilmente que la *ley producto* es asociativa, que admite por elemento neutro $e = (e_1, e_2)$, e_1 y e_2 son los elementos neutros respectivos de G_1 y G_2; en fin, todo elemento (x_1, x_2) es inversible y

$$(x_1, x_2)^{-1} = (x_1^{-1}, x_2^{-1}),$$

$G_1 \times G_2$ provisto de la ley definida por (3) (o (1) o (2)) tiene, pues, una estructura de grupo, se le llama *grupo producto* (cartesiano) de los grupos G_1 y G_2.

$G_1 \times G_2$ es conmutativo si y sólo si G_1 y G_2 lo son.

Se podrá definir de la misma manera $G_1 \times G_2 \times \ldots \times G_n$ *producto cartesiano de n grupos.* Se podrá considerar en particular el caso en que $G_1 = G_2 = \ldots = G_n = G$; por ejemplo, $G \times G$ (si se emplea la notación G^2, no hay que confundir este grupo con el conjunto descrito por *xy*, *x* e *y* variando en G (ver § 50)). Por ejemplo, si G es un grupo aditivo de elemento neutro 0, en G^n (en el sentido precedente), tendremos

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$
$$e = (0, 0, \ldots, 0)$$
$$-(x_1, x_2, \ldots, x_n) = (-x_1, -x_2, \ldots, -x_n).$$

Volvamos al grupo producto $G_1 \times G_2$ de dos grupos cualesquiera, si H_1 es un subgrupo de G_1 y H_2 de G_2 se ve inmediatamente que $H_1 \times H_2$ es un subgrupo de $G_1 \times G_2$; en efecto, para todo par (x_1, x_2) e (y_1, y_2) de $H_1 \times H_2$, tenemos

$$(x_1, x_2)(y_1, y_2)^{-1} = (x_1 y_1^{-1}, x_2 y_2^{-1}) \in H_1 \times H_2,$$

pues $x\,y^{-1}$ pertenece a H y $x\,y^{-1}$ a H .

EJERCICIOS

1. Demostrar que $\mathbf{Z}^n$ es un subgrupo de $\mathbf{R}^n$ (para la adición) y que $\mathbf{R}^n/\mathbf{Z}^n$ es isomorfo a $\mathbf{T}^n$ («toro de *n* dimensiones», ver § 77, ej. 3).
2. Siendo H_1 y H_2 dos subgrupos distinguidos respectivos de G_1 y G_2, demostrar que $H_1 \times H_2$ es un subgrupo distinguido de $G_1 \times G_2$ y que $(G_1 \times G_2)/(H_1 \times H_2)$ es isomorfo a $(G_1/H_1) \times (G_2/H_2)$.

80. Relaciones entre $G_1 \times G_2$ y ciertos de sus subgrupos

a) Siendo G_1 y $\{e_1\}$ subgrupos de G_1 y G_2 y $\{e_2\}$ de G_2, $G_1 \times \{e_2\}$ y $\{e_1\} \times G_2$ son subgrupos de $G_1 \times G_2$. Consideremos la aplicación

$$f\colon\ G_1 \to G_1 \times \{e_2\}$$

definida por $f(x_1) = (x_1, e_2)$, es manifiestamente biyectiva; por otra parte,

$$f(x_1 y_1) = (x_1 y_1, e_2) = (x_1, e_2)(y_1, e_2) = f(x_1) f(y_1)$$

luego es un isomorfismo, de donde:

TEOREMA. — *Los grupos* G_1 *y* G_2 *son, respectivamente, isomorfos a los subgrupos*

$$G'_1 = G_1 \times \{e_2\} \quad y \quad G'_2 = \{e_1\} \times G_2 \quad de \quad G_1 \times G_2.$$

Sucede, como ya lo hemos hecho, que se identifica G_1 y $G'_1 = G_1 \times \{e_2\}$ (o G'_1 y G'_2) *por esta identificación* G_1 *y* G_2 *pueden ser consideradas como subgrupos de* $G_1 \times G_2$.

Por ejemplo, en el plano $\mathbf{R}^2$, esto equivale a confundir el *punto x de la recta* O*x* y el *punto* (*x*, 0) *del plano.*

b) Consideremos la primera proyección pr_1 de $G_1 \times G_2$ sobre G_1 (ver § 12, *d*)

$$(x_1, x_2) \to pr_1(x_1, x_2) = x_1$$

tenemos

$$pr_1[(x_1, x_2)(y_1, y_2)] = pr_1(x_1y_1, x_2y_2) = x_1y_1 = pr_1(x_1, x_2)pr_1(y_1, y_2),$$

luego es un *homomorfismo*. Busquemos su núcleo, es decir, $pr_1^{-1}(e_1)$ es el conjunto de los elementos (x_1, x_2) para los cuales $x_1 = e_1$, es decir,

$$pr_1^{-1}(e_1) = \{e_1\} \times G_2 = G_2'.$$

G_2' es, pues, un subgrupo distinguido de $G_1 \times G_2$ (ver § 77, t. 2). Consideremos el grupo cociente $(G_1 \times G_2)/G_2'$, la relación de equivalencia módulo G_2' es equivalente a

$$pr_1(x_1, x_2) = pr_1(y_1, y_2) \Leftrightarrow x_1 = y_1;$$

dicho de otra manera, una clase $\left(\widehat{x_1, x_2}\right)$ módulo G_2' está descrita por los elementos (x_1, x_2) con x_1 fijo, la aplicación

$$G_1 \to (G_1 \times G_2)/G_2'$$

definida por $x_1 \to \left(\widehat{x_1, x_2}\right)$ es una biyección, es un homomorfismo, como se ve inmediatamente; es, pues, un isomorfismo, de donde:

TEOREMA. — *La proyección de $G_1 \times G_2$ sobre G_1 (resp. G_2) es un homomorfismo; su núcleo es $G_2' = \{e_1\} \times G_2$ (resp. $G_1' = G_1 \times \{e_1\}$ y el grupo cociente $(G_1 \times G_2)/G_2'$ (resp. $(G_1 \times G_2)/G_1'$) es isomorfo a G_1 (resp. a G_2).*

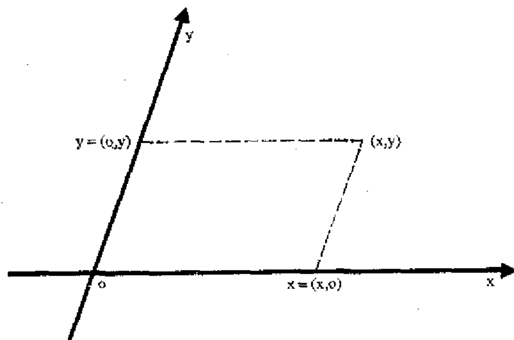


FIG. 10

81. Suma directa (en el caso de grupos conmutativos)

a) Reiniciemos el estudio precedente y supongamos G_1 y G_2 abelianos, estando la ley representada aditivamente, tenemos

(1) $$x = (x_1, x_2) = (x_1, e_2) + (e_1, x_2),$$

es decir (ver § 50),

$$G_1 \times G_2 = G_1' + G_2'$$

y la descomposición de un elemento cualquiera x de $G_1 \times G_2$ se efectúa de una *manera única* en suma de un elemento de G_1' y de un elemento de G_2'; por otra parte, si (x_1, x_2) pertenece a G_1' y a G_2', $x_1 = e_1$ y $x_2 = e_2$, luego poniendo $e = (e_1, e_2)$, elemento neutro de $G_1 \times G_2$,

$$(2) \qquad G_1' \cap G_2' = \{e\}.$$

b) De una manera general, sea G un grupo abeliano (con notación aditiva) que es la suma de dos de sus subgrupos H_1 y H_2

$$G = H_1 + H_2,$$

supongamos, además, que la descomposición $x = x_1 + x_2$ (x_1 elemento de H_1, x_2 de H_2) sea *única* para todo x de G y busquemos en este caso los elementos comunes a H_1 y H_2, tendremos

$$x \in H_1 \cap H_2 \qquad \begin{matrix} x = x + e & x \in H_1 & e \in H_2 \\ x = e + x & e \in H_1 & x \in H_2 \end{matrix}$$

de donde, según la unicidad de la descomposición, $x = e$.

Recíprocamente, supongamos $G = H_1 + H_2$ y $H_1 \cap H_2 = \{e\}$ si

$$\begin{matrix} x = x_1 + x_2 = y_1 + y_2 & (x_1, y_1 \in H_1, x_2, y_2 \in H_2) \\ x_1 - y_1 = x_2 - y_2 & (x_1 - y_1 \in H_1, x_2 - y_2 \in H_2), \end{matrix}$$

luego $x_1 - y_1 = x_2 - y_2 = e$ y la descomposición es única, de donde:

TEOREMA Y DEFINICIÓN. — *Para un grupo abeliano* G *que es suma de dos de sus subgrupos* H_1, H_2, *las dos propiedades siguientes son equivalentes:*

a) *Todo x de* G *se escribe de una manera única*

$$x = x_1 + x_2 \qquad (x_1 \in H_1, x_2 \in H_2).$$

b) $H_1 \cap H_2 = \{e\}$.

Se dice entonces que G *es* suma directa *de los subgrupos* H_1 *y* H_2 *y se escribe*

$$G = H_1 \oplus H_2.$$

Consideremos la aplicación f de $H_1 \times H_2$ en G definida por

$$(x_1, x_2) \rightarrow x_1 + x_2 = f(x_1, x_2).$$

Es evidentemente biyectiva, pues G es suma directa de H_1 y H_2, luego todo x de G corresponde a una pareja única (x_1, x_2); por otra parte,

$$\begin{aligned} f[(x_1, x_2) + (y_1, y_2)] &= f(x_1 + y_1, x_2 + y_2) = (x_1 + y_1) + (x_2 + y_2) \\ &= (x_1 + x_2) + (y_1 + y_2) = f(x_1, x_2) + f(y_1, y_2), \end{aligned}$$

luego, acudiendo a un resultado del párrafo 80:

TEOREMA. — $G = H_1 \oplus H_2$ *es isomorfo al grupo producto* $H_1 \times H_2$. *Y* G/H_1 *es isomorfo a* H_2 (G/H_2 *es isomorfo a* H_1).

OBSERVACION

Se puede intentar extender esta teoría a los grupos no conmutativos (representados multiplicativamente). Siendo H_1 y H_2 dos subgrupos cualesquiera de G, hay que observar que, si $H_1 \times H_2$ es siempre un subgrupo de G, no lo es en general de H_1H_2 (ver ej. 1); la teoría es, pues, menos simple, es el objeto de los ejercicios 2 y 3 siguientes.

EJERCICIOS

1. Siendo H_1 y H_2 dos subgrupos invariantes de G, mostrar que H_1H_2 es un subgrupo de G (que es invariante). Demostrar que este resultado es falso para dos subgrupos cualesquiera (tomar en el grupo de las rotaciones del espacio alrededor de un punto O $H_1 = \{I, S_1\}$ $H_2 = \{I, S_2\}$, I identidad, S_1 y S_2 rotaciones de ángulo π alrededor de OD_1 y OD_2 distintas y no perpendiculares).
2. Se toman las notaciones del párrafo. Demostrar que todo (x_1, x_2) de G se escribe de una manera única como producto de (x_1, e_2) de G_1' y (e_1, x_2) de G_2', que todo elemento de G_1' permuta con todo elemento de G_2' y finalmente que el único elemento común a G_1' y G_2' es (e_1, e_2).
3. Sea G un grupo y dos de s subgrupos H_1 y H_2 tales que $G = H_1H_2$. Demostrar que las condiciones siguientes *a*) y *b*) son equivalentes:

a) Para todo x de G, existe x_1 de H_1 *único* y x_2 de H_2 *único* tales que $x = x_1x_2$. Todo elemento de H_1 permuta con todo elemento de H_2.

b) H_1 y H_2 son subgrupos invariantes y $H_1 \cap H_2 = \{e\}$. Se dice entonces que G es *producto directo* de H_1 y H_2. Demostrar que $G = H_1H_2$ es en este caso isomorfo a $H_1 \times H_2$ (en este caso no hay inconveniente en confundir producto H_1H_2 y producto cartesiano $H_1 \times H_2$).

4. Dado un grupo abeliano G representado aditivamente suma de n de sus subgrupos $H_1, \ldots, H_n$, se dice que G es *suma directa* de $H_1, \ldots, H_n$ y se escribe

$$G = H_1 \oplus H_2 \oplus \ldots \oplus H_n$$

si y sólo si para todo x existe x_i de H_i (para todo $i \in [1, n]$ *único* tal que $x = x_1 + x_2 + \ldots + x_n$. Demostrar que la condición precedente es equivalente a la siguiente: para todo $i \in [1, n-1]$

$$(H_1 + H_2 + \ldots + H_i) \cap H_{i+1} = \{e\}.$$

V. Generación de grupos

82. Parte generatriz de un grupo

Hemos visto la definición del subgrupo de G engendrado por una parte A de G (ver § 73); puede suceder que el subgrupo engendrado por A sea el mismo G, se dice entonces que A es una *parte generadora*[15] de G. Si $A = (a_i)_{i \in I}$, se dice que (a_i) es una *familia generadora* de G.

(15) Preferimos las expresiones «*parte generadora*» y «*familia generadora*» a la expresión tradicional «*sistema de generadores*» que podría hacer creer que *todo* elemento de A es un generador de G, lo que no es cierto, salvo que A contenga un solo elemento.

Si G está engendrado por una de sus partes *finitas,* se dice que G es de *tipo finito* (no hay que confundirlo con un grupo de *orden finito*). Por ejemplo, un grupo engendrado por un elemento único a se llama *monógeno,* este grupo (representado multiplicativamente) está descrito por a^n, n describiendo **Z**. El mismo **Z** es un grupo monógeno engendrado por 1.

EJERCICIOS

1. Sea P la *recta proyectiva real* (es decir, **R** completado por un punto al infinito de manera que $x \mapsto 1/x$ sea una biyección de P). Se consideran las aplicaciones f_1 f_2 de P sobre sí mismo definidas por $f_1(x) = 1/x$, $f_2(x) = 1 - x$.

Demostrar que el grupo G (para la composición de aplicaciones) engendrado por f_1 y f_2 es de orden 6. Formar la tabla de este grupo.

2. Se considera el grupo engendrado por dos elementos a y b tales que (n, entero > 0)

$$a^n = e \qquad b^2 = e \qquad aba = b.$$

a) Formar la tabla de este grupo para $n = 2$ y $n = 4$.

b) Demostrar que el grupo de las isometrías conservando un polígono regular de n lados está engendrado por dos isometrías a y b que verifican para la composición de apli caciones las relaciones precedentes.

3. Demostrar que $\mathbf{Z}^n$ es un grupo de tipo finito. Se indicará una familia generatriz de n elementos.

83. Grupo monógeno. Grupo cíclico

a) Consideremos un grupo *monógeno,* es decir, engendrado por un elemento a; en notación multiplicativa

$$G = \{\ldots\ a^{-p},\ \ldots,\ a^{-1},\ e,\ a,\ \ldots,\ a^p\ \ldots\}.$$

Este grupo es abeliano. Sea, por otra parte, la aplicación f del grupo aditivo **Z** en el grupo G definida por

$$p \longrightarrow f(p) = a^p.$$

Se tiene $f(p + q) = a^{p+q} = a^p a^q = f(p)f(q)$, luego f es un *homomorfismo,* según la definición de G, f es *suprayectiva.* Su núcleo $f^{-1}(e)$ es un subgrupo de **Z**, sea $n\mathbf{Z}$ ($n \geq 0$); luego $\mathbf{Z}/n\mathbf{Z}$ es isomorfo a $f(\mathbf{Z}) = G$ (ver § 78). Hay que considerar dos casos:

1.° $n = 0$, $n\mathbf{Z} = \{0\}$ y $\mathbf{Z}/\{0\} = \mathbf{Z}$.

2.° $n > 0$, $n\mathbf{Z} \neq \{0\}$ y $\mathbf{Z}/n\mathbf{Z} = \left(\dot{0},\ \dot{1},\ \ldots,\ \widehat{n-1}\right)$.

En este último caso

$$G = \{a^0 = e,\ a,\ a^1,\ \ldots,\ a^{n-1}\}.$$

Se dice entonces que G de orden n es un *grupo cíclico de orden n.*

Este estudio se puede resumir en el siguiente teorema:

TEOREMA. — *Todo grupo monógeno es isomorfo:*

— *A* **Z** *si es infinito.*
— *A* **Z**/*n***Z** *si es de orden n.*

b) En un grupo cualquiera, se dice que un elemento a es de *orden p* si el subgrupo engendrado por a es de orden p: este subgrupo es entonces un *grupo cíclico de orden p.*

En particular, si G es un grupo finito de orden n, todos los elementos de G son de orden finito según el teorema del § 74, *b*), el orden p de un subgrupo engendrado por un elemento cualquiera es un divisor de n, luego $n = pq$ y

$$a^n = a^{pq} = (a^p)^q = e^q = e.$$

TEOREMA. — *En un grupo finito* G *de orden n, el orden de cada uno de los elementos es un divisor de n y para todo elemento de* G: $a^n = e$.

EJERCICIOS

1. Demostrar que el grupo de las rotaciones del plano de centro O y de ángulo $k2\pi/n$ (n entero > 0 fijado, k entero racional cualquiera) es un grupo cíclico de orden n.
2. Demostrar directamente (sin referencia al teorema del § 74) que el orden p de todo elemento de un grupo cíclico de orden n es un divisor de n.
3. Colocar según su orden los elementos de un grupo cíclico de orden 30.
4. Demostrar que todo subgrupo de un grupo cíclico es cíclico. ¿Cuáles son los subgrupos de un grupo cíclico de orden 30 (ver ej. 3)?
5. ¿En qué condición el elemento p de $\mathbf{Z}/n\mathbf{Z}$ engendra el grupo aditivo $\mathbf{Z}/n\mathbf{Z}$?
6. Demostrar que todo grupo finito de orden p (p entero primo) es cíclico y que está engendrado por cada uno de sus elementos distintos del elemento neutro.

VI. Grupos de transformaciones

84. Grupo de permutaciones de un conjunto E

Se llama *permutación* de E *toda biyección* de E sobre sí mismo (ver § 40, *c*); el conjunto de todas las permutaciones de E es un grupo para la composición de aplicaciones:

1.ª $(f, g) \to f \circ g$ es una ley interna asociativa (la compuesta de dos biyecciones es una biyección).
2.ª id_E es una permutación de E (permutación idéntica).
3.ª Si f es una permutación de E, f^{-1} también lo es.

Este grupo se llama el *grupo de las permutaciones de* E, o *grupo simétrico* de E; se le presenta por $\mathfrak{S}_E$.

Consideremos dos conjuntos E y E′ *equipotentes,* existe entonces una biyección φ de E sobre E′; consideremos la aplicación de $\mathfrak{S}_E$ en $\mathfrak{S}_{E'}$ definida por

$$(1) \qquad f \to \varphi \circ f \circ \varphi^{-1} = f'$$

cualquiera que sea f' de $\mathcal{S}_{E'}$, tendremos

$$\varphi \circ f \circ \varphi^{-1} = f' \Leftrightarrow f = \varphi^{-1} \circ f' \circ \varphi$$

la aplicación considerada de $\mathcal{S}_E$ en $\mathcal{S}_{E'}$ es, pues, una *biyección*; por otra parte,

$$f' \circ g' = (\varphi \circ f \circ \varphi^{-1}) \circ (\varphi \circ g \circ \varphi^{-1}) = \varphi \circ (f \circ g) \circ \varphi^{-1}.$$

De donde: *Los grupos $\mathcal{S}_E$ y $\mathcal{S}_{E'}$ de dos conjuntos equipotentes son isomorfos para la composición de aplicaciones.*

85. Estudio del grupo simétrico $\mathcal{S}_n$. Notaciones diversas

Supongamos card $E = n$, todos los grupos $\mathcal{S}_E$ son isomorfos a $\mathcal{S}_n$, grupo de las permutaciones de $[1, n]$. Todos estos grupos son, pues, de orden $n!$ (ver § 40). Estudiaremos las permutaciones de $[1, n]$; en los ejemplos en que n está especificado, por ejemplo, para $n = 5$, consideramos también los conjuntos tales como

$$E = \{a, b, c, d, e\}.$$

Una permutación p de $[1, n]$ se representará $i \to p(i) = p_i$ o también

$$\begin{pmatrix} 1 & 2 & \dots & i & \dots & n \\ p_1 & p_2 & \dots & p_i & \dots & p_n \end{pmatrix}$$

esta última notación es evidente, o también si $i \to a_i$ es una biyección de $[1, n]$

$$\begin{pmatrix} a_1 & a_2 & \dots & a_i & \dots & a_n \\ p_{a_1} & p_{a_2} & \dots & p_{a_i} & \dots & p_{a_n} \end{pmatrix}.$$

Por abuso de lenguaje diremos que $p \circ q$ o pq es el "*producto*" de dos permutaciones p y q (ver § 15, nota 1).

Conforme a las notaciones indicadas § 70, escribiremos

$$p^2 = p \circ p \qquad p^k = p^{k-1} \circ p.$$

El grupo $\mathcal{S}_n$ no es conmutativo para $n \geq 3$; supongamos que p y q dejan invariantes todos los elementos de $[1, n] - \{a, b, c\}$

$$p = \begin{pmatrix} a & b & c & x & \dots & y \\ b & c & a & x & \dots & y \end{pmatrix} \qquad q = \begin{pmatrix} a & b & c & x & \dots & y \\ b & a & c & x & \dots & y \end{pmatrix},$$

tendremos

$$pq = \begin{pmatrix} a & b & c & x & \dots & y \\ c & b & a & x & \dots & y \end{pmatrix} \neq qp = \begin{pmatrix} a & b & c & x & \dots & y \\ a & c & b & x & \dots & y \end{pmatrix}.$$

(Recordemos que tanto para $p \circ q$ como para pq la permutación q se efectúa la primera.)

Diremos que una permutación p *opera sobre una parte* P *de* E *si* p *deja invariante cada elemento de* E — P. Se ve inmediatamente que si p y q operan, respectivamente, sobre dos partes P y Q *disjuntas*: $pq = qp$.

Se llama *transposición* una permutación t tal que $(i \neq j)$

$$k \neq i \text{ y } k \neq j \qquad t(a_k) = a_k$$
$$t(a_j) = a_i \qquad t(a_i) = a_j,$$

es decir, es una operación operando sobre $\{a_i, a_j\}$ y distinta de la identidad.

De una manera más general, se llamará *ciclo* una permutación c tal que siendo p un entero natural $(1 \leq p \leq n)$, se tenga

$$(i = 1, 2, \ldots, p-1) \qquad c(a_i) = a_{i+1}$$
$$c(a_p) = a_1$$
$$(j = p+1, \ldots, n) \qquad c(a_j) = a_j$$

se representará, pues, un tal ciclo c por

$$c = \begin{pmatrix} a_1 & a_2 & \ldots & a_{p-1} & a_p & a_{p+1} & \ldots & a_n \\ a_2 & a_3 & \ldots & a_p & a_1 & a_{p+1} & \ldots & a_n \end{pmatrix}$$

o más simplemente

$$c = (a_1, a_2, \ldots, a_p),$$

siendo cada elemento no indicado invariante por c, cada uno de los elementos indicados, salvo a_p, tiene por imagen por c el siguiente y a_p tiene por imagen a_1. Vemos que

$$c^2 = \begin{pmatrix} a_1 & a_2 & \ldots & a_p & a_{p+1} & \ldots & a_n \\ a_3 & a_4 & \ldots & a_2 & a_{p+1} & \ldots & a_n \end{pmatrix}$$

y para $1 < k < p$

$$c^k = \begin{pmatrix} a_1 & \ldots\ldots\ldots & a_p & a_{p+1} & \ldots & a_n \\ a_{k+1} & \ldots\ldots\ldots & a_k & a_{p+1} & \ldots & a_n \end{pmatrix} \neq u$$

$$c^p = \begin{pmatrix} a_1 & \ldots\ldots\ldots & a_p & a_{p+1} & \ldots & a_n \\ a_1 & \ldots\ldots\ldots & a_p & a_{p+1} & \ldots & a_n \end{pmatrix} = u$$

designando por u la *permutación idéntica*, resulta $c^p = u$ y p es el menor entero $h > 0$ tal que $c^h = u$.

Luego el ciclo $c = (a_1, a_2, \ldots, a_p)$ *operando sobre* p *elementos de un conjunto de* n *elementos es un elemento de orden* p *del grupo* $\mathcal{S}_n$.

(Se verifica, pues, en este caso particular que el orden p divide el orden de grupo $\mathcal{S}_n$ sea $n!$, puesto que $1 \leq p \leq n$.)

Un ciclo de orden p *se llama también una permutación circular de orden* p.

Una transposición es, pues, un ciclo de orden 2; la permutación idéntica u es un ciclo de orden 1 (es, por otra parte, el único).

Según una observación hecha más arriba, dos ciclos c_1 y c_2 operando sobre dos *partes disjuntas* de E, son *permutables*.

86. Descomposición de una permutación en producto de ciclos

Sea p una permutación de $E = [1, n]$.

Si $p \neq u$, existe un elemento de E sea a_1 tal que $p(a_1) \neq a_1$; pongamos $p(a_1) = a_2$, $p(a_2) = a_3$ y sea i el primer índice tal que

$$p(a_i) \in \{a_1, a_2, \ldots, a_i\}.$$

Veamos que $p(a_i) = a_1$; en efecto, si $p(a_i) = a_{i'}$ $(1 < i' \leq i - 1)$, se tendría $p(a_i) = p(a_{i'-1})$, lo que es imposible, pues si a_i y $a_{i'-1}$, distintos, tendrían igual imagen. Luego la *restricción* de p a $\{a_1, a_2, \ldots, a_i\}$ es un *ciclo de orden i.*

Designemos por a_{i+1} un elemento de $E - \{a_1, a_2, \ldots, a_i\}$; pondremos en evidencia un nuevo ciclo de orden j operando sobre $\{a_{i+1}, \ldots, a_{i+j}\}$. Siendo E finito, al cabo de un número finito de operaciones habremos agotado E; habremos así definido una partición de E de elementos P_i, P_j, ..., P_l, tal que la restricción de p a cada uno de los elementos de la partición sea un ciclo, el orden del ciclo operando sobre P_k es el número de elementos de P_k; de donde:

Toda permutación de un conjunto de n elementos se puede descomponer en producto de ciclos cuya suma de órdenes es n, siendo permutables dos ciclos cualesquiera.

EJEMPLO

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 1 & 4 & 2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 2 & 6 & 5 \\ 6 & 5 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$
$$= (1, 3)\ (2, 6, 5)\ (4) = (1, 3)\ (2, 6, 5),$$

pues todo ciclo de orden 1 es idéntico a u (naturalmente hay que contarlo en la descomposición si se quiere que la suma de los órdenes de los ciclos sea igual a n).

Por otra parte, todo ciclo se puede descomponer en producto de transposición; en efecto,

$$(a_1, a_2, \ldots, a_p) = (a_1, a_2)(a_2, a_3) \ldots (a_{p-2}, a_{p-1})(a_{p-1}, a_p).$$

(ATENCIÓN: Hay que empezar por la derecha). Por ejemplo,

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 4 & 2 \\ 2 & 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 4 & 3 \\ 1 & 3 & 4 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \end{pmatrix}$$
$$= (1, 2)\ (2, 3)\ (3, 4).$$

Pero en este caso dos de las transposiciones (a_1, a_2), ..., (a_{p-1}, a_p) *en general no son permutables.*

Estos dos resultados nos permiten decir:

Toda permutación de un conjunto de n elementos es descomponible en producto de transposiciones.

EJERCICIOS

1. ¿Cuántas transposiciones de ciclos de orden 3, de ciclos de orden p hay en $\mathcal{S}_n$?

2. Una permutación que es descomponible en r ciclos de órdenes respectivos $k_1, k_2, \ldots, k_r$, ¿en cuántas transposiciones se puede descomponer?

3. Demostrar que el grupo simétrico $\mathcal{S}_n$ está engendrado por las $n-1$ transposiciones

$$(1, 2), \quad (2, 3), \quad \ldots, \quad (n-1, n).$$

Se demostrará primero que toda transposición (i, j) —con $i<j$— es descomponible en producto de transposición $(k, k+1)$, $1 \leq k \leq n-1$.

Demostrar que al quitar una de estas transposiciones, por ejemplo, $(p, p+1)$, las transposiciones que quedan

$$(1, 2) \;\; (2, 3) \;\; \ldots \;\; (p-1, p) \;\; (p+1, p+2) \;\; \ldots \;\; (n-1, n)$$

operan sobre dos partes disjuntas de $[1, n]$ y, por consiguiente, no pueden engendrar $\mathcal{S}_n$, es decir, que las $n-1$ transposiciones dadas más arriba forman una familia generatriz minimal de $\mathcal{S}_n$.

4. Se pone $t=(1, 2)$ y $c=(1, 2, \ldots, n)$, calcular c^k, a continuación $c^k t c^{-k}$ $(1 \leq k \leq n)$, deducir de lo anterior que t y c engendran $\mathcal{S}_n$.

87. Signatura de una permutación

Sea p una permutación de $[1, n]$; siendo p biyectiva

$$i<j \Rightarrow \text{ o bien } p(i)<p(j) \text{ o bien } p(i)>p(j)$$

en el segundo caso diremos que los dos elementos $p(i)$ y $p(j)$ presentan una *inversión.* Designemos por $I(p)$ el *número total de inversiones* presentadas dos a dos por los elementos de $\{p(1), \ldots, p(n)\}$ y consideremos el producto

$$V_n = \prod_{1 \leq i<j \leq n} (j-i) = [2-1][(3-1)(3-2)]\ldots[(n-1)(n-2)\ldots(n-(n-1))].$$

Este producto es un entero estrictamente positivo; escribamos igualmente

$$p(V_n) = \prod_{1 \leq i<j \leq n} [p(j)-p(i)],$$

donde p es una biyección, cada factor de V_n se vuelve a encontrar en $p(V_n)$ una vez y sólo una vez, salvo el signo, se ve que

$$p(V_n) = (-1)^{I(p)} V_n,$$

la aplicación de $\mathfrak{S}_n$ en $\{1, -1\}$ definida por

$$p \to \varepsilon(p) = (-1)^{I(p)}$$

se llama la *signatura de la permutación p.*

Si $\varepsilon(p) = +1$ se dice que p es una *permutación par.*

Si $\varepsilon(p) = -1$ se dice que p es una *permutación impar.*

Luego

$$p(V_n) = \varepsilon(p)V_n.$$

Naturalmente $\varepsilon(u) = 1$. Busquemos la paridad de una transposición, se la puede escribir (i, j) con $i < j$ o mejor

$$t = \begin{pmatrix} 1 & 2 & \dots & i-1 & i & i+1 & \dots & j-1 & j & j+1 & \dots & n \\ 1 & 2 & \dots & i-1 & j & i+1 & \dots & j-1 & i & j+1 & \dots & n \end{pmatrix}$$

busquemos el número total de inversiones de t. Los elementos de $[1, n] - \{i, j\}$ no presentan ninguna inversión dos a dos.

$p(i) = j$ y $p(j) = i$ no presentan ninguna inversión con los elementos de $[1, i-1]$ y de $[j+1, n]$, pero $p(i)$ y $p(j)$ presentan una inversión y

$p(i)$ presenta una inversión con cada elemento de $[i+1, j-1]$,
$p(j)$ presenta una inversión con cada elemento de $[i+1, j-1]$,

luego el número total de inversiones de una transposición es impar:

Toda transposición es impar.

Consideremos dos permutaciones p y q; busquemos la signatura de $p \circ q$, tenemos

$$(p \circ q)(V_n) = p[q(V_n)] = \varepsilon(p)q(V_n) = \varepsilon(p)\varepsilon(q)V_n,$$

pero

$$(p \circ q)V_n = \varepsilon(p \circ q),$$

luego

$$\varepsilon(p \circ q) = \varepsilon(p)\varepsilon(q).$$

Es decir, *la aplicación $p \to \varepsilon(p)$ es un homomorfismo suprayectivo del grupo simétrico $\mathfrak{S}_n$ sobre el grupo multiplicativo $\{1, -1\}$.*

Se deduce que *el producto de dos permutaciones de la misma paridad es una permutación par y que p y p^{-1} tienen la misma paridad*; por consiguiente, resulta que *el conjunto de las permutaciones pares es un subgrupo de $\mathfrak{S}_n$*, se le llama el *grupo alternado* y se le representa por $\mathfrak{A}_n$.

En fin, si una transposición es impar, toda permutación *par* (resp. *impar*) es descomponible en producto de un número *par* (resp. *impar*) de *transposiciones.*

EJERCICIOS

1. Determinar las permutaciones pares y las permutaciones impares de $\mathcal{S}_3$.
2. ¿Cuál es la paridad del ciclo $C = (1, 2, ..., p)$?
3. Demostrar que $\mathcal{A}_n$ es un *subgrupo invariante* de $\mathcal{S}_n$ (se observará que $\mathcal{A}_n$ es el núcleo del homomorfismo $p \to \varepsilon(p)$). Deducir que $\mathcal{A}_n$ es de orden $n!/2$ (considerar el grupo cociente $\mathcal{S}_n/\mathcal{A}_n$).

88. Grupo de transformaciones operando en un conjunto

a) *Se llama grupo de transformaciones de* E, *o grupo de permutaciones de* E *todo subgrupo* G *del grupo simétrico* $\mathcal{S}_E$. Se dice que G *opera* en E.

ATENCIÓN: No se debe confundir *el* grupo *de las* permutaciones de E ($\mathcal{S}_E$) y *un* grupo *de* permutaciones de E (G subgrupo de $\mathcal{S}_E$).

EJEMPLOS Y EJERCICIOS

1. El grupo de las traslaciones, el grupo de los desplazamientos, el grupo de las isometrías del plano (o del espacio) son grupos de transformaciones del plano (o del espacio).
2. El conjunto de las rotaciones planas (o el conjunto de las semejanzas planas) de centro O fijo es un grupo de transformaciones del plano.
3. Siendo G un grupo de transformaciones de E, el conjunto H de las aplicaciones de G dejando invariante un elemento a (o una parte A) de E es un subgrupo de G.
4. El grupo G′ de los automorfismos interiores de un grupo G (ver § 77, ej. 4) es un grupo de transformaciones de G.

b) Sea G′ un grupo cualquiera y φ un isomorfismo de G′ sobre un grupo G de transformaciones de un conjunto E, se dice que el grupo G es una *realización del grupo* G′ *como grupo de transformaciones de* E.

Consideremos un grupo cualquiera G y el conjunto de las traslaciones por la izquierda Γ; consideremos la aplicación φ de G sobre Γ definida por

$$a \to \varphi(a) = \gamma_a$$

φ es suprayectiva, es inyectiva, pues,

$$\gamma_a = \gamma_{a'} \Leftrightarrow (\forall\, x \in G) \qquad ax = a'x \Leftrightarrow a = a',$$

pues en un grupo todo elemento es regular.

Por otra parte,

$$(\forall\, x \in G) \qquad (\gamma_a \circ \gamma_b)(x) = \gamma_a[\gamma_b(x)] = \gamma_a(bx) = a(bx) = (ab)x = \gamma_{ab}(x),$$

luego para todo par (a, b)

$$\gamma_{ab} = \gamma_a \circ \gamma_b.$$

Luego φ *es un isomorfismo del grupo* G *y de* Γ *provisto de la composición de aplicaciones*; Γ *es, pues, un grupo,* podemos entonces decir:

Todo grupo G *puede ser realizado como grupo de transformaciones de sí mismo.*

OBSERVACION

Para esta demostración se hubiera podido utilizar el grupo Δ de las traslaciones por la derecha, pero en este caso Δ es isomorfo a G_1, conjunto G provisto de la ley $(a, b) \to ba$ (*ley opuesta* a la ley de G, ver § 44); se demostrará a título de ejercicio que G_1 es un grupo isomorfo a G (se le llama *grupo opuesto* a G).

49. Transitividad e intransitividad

a) Un grupo G de transformaciones de E es *transitivo* si, cualesquiera que sean los dos elementos *a* y *b* de E, existe una permutación *f* de G tal que $b = f(a)$; en el caso contrario el grupo se llama *intransitivo.*

Cuando G es transitivo se dice que *opera transitivamente* en E, el *conjunto* E *provisto de* G se llama entonces *espacio homogéneo.*

EJEMPLOS

1. El grupo de traslación del plano (o del espacio) es transitivo; luego el plano (o el espacio) provisto del grupo de las traslaciones es un espacio homogéneo.
2. El grupo de las rotaciones del plano de centro fijo O es intransitivo.

b) Dado un grupo de transformaciones operando en E, consideremos la relación siguiente definida en E

$$I(a, b) \Leftrightarrow (\exists f \in G) \qquad b = f(a)$$

es una relación de equivalencia:

— I es *reflexiva*

$$(\forall a \in E) \qquad (\exists u \in G) \qquad u(a) = a.$$

— I es *simétrica*

$$[(\exists f \in G) \quad b = f(a)] \Rightarrow [(\exists f^{-1} \in G) \quad a = f^{-1}(b)].$$

— I es *transitiva*

$$\left\{\begin{array}{ll} \exists f \in G & b = f(a) \\ \exists g \in G & c = g(b) \end{array}\right. \Rightarrow [(\exists g \circ f \in G) \quad c = (g \circ f)(a)].$$

La relación I es, pues, una *relación de equivalencia definida sobre* E, sus clases se llaman las *clases de intransitividad* del grupo G; la clase $\dot{a}$ de *a* no es otra cosa que la parte de E descrita por $f(a)$; cuando *f* describe G, se dice entonces que esta clase es la *órbita* de *a* para el grupo G.

Se ve que G es transitivo si y sólo si todos los elementos de E son equivalentes módulo I, es decir, si la relación I es la equivalencia absoluta.

EJEMPLOS Y EJERCICIOS

1. Encontrar las clases de intransitividad para los conjuntos E provistos de los grupos de transformaciones G siguientes:

a)	E plano (o espacio)	G: Translaciones paralelas a una dirección de recta.
b)	E plano	G: Rotaciones de centro O.
	E plano	G: Rotaciones de centro O, de ángulo $k2\pi/n$ ($n > 0$, k entero cualquiera).
c)	E espacio	G: Rotaciones de eje fijo Δ.
	E espacio	G: Rotaciones de eje pasando por O fijo.

2. Sea G un grupo y G′ el grupo de los automorfismos interiores de G (§ 77, *a*) y ej. 4); la relación x y x' son conjugadas en G (§ 77, ej. 7) no es otra que la equivalencia «existe f de G′ tal que $x' = f(x)$».

Ejercicios

60. Para cada uno de los subconjuntos siguientes X del plano o del espacio:

a) triángulo equilátero b) cuadrado
c) rectángulo no cuadrado d) tetraedro regular
e) tres semirrectas Ox, Oy, Oz ortogonales dos a dos
f) tres ejes x'Ox, y'Oy, z'Oz ortogonales dos a dos
g) tres rectas, pasando por O ortogonales dos a dos

determinar el conjunto I de las isometrías y el conjunto D de los desplazamientos (§ 70, ej. 6) para los que X es invariante.

En cada caso demostrar que para la composición de aplicaciones I es un grupo y D uno de los subgrupos.

Determinar en cada caso los subgrupos de I y los subgrupos de D. (Se podrá escribir ab en lugar de $a \circ b$ para simplificar la escritura).

61. Se considera el conjunto de las biyecciones siguientes, operando en la *recta proyectiva real* $\tilde{\mathbf{R}}$ (obtenida adjuntando a **R** un punto al infinito $\infty = 1/0$) o en el *plano analítico* $\tilde{\mathbf{C}}$ (obtenido añadiendo a **C** un punto al infinito $\infty = 1/0$, ver capítulo 6, § 124)

$$x \mapsto x, \qquad x \mapsto 1/x, \qquad x \to -x, \qquad x \to -1/x.$$

Demostrar que para la composición de aplicaciones este conjunto es un grupo isomorfo al grupo de las isometrías de un rectángulo (ej. 60).

62. Se considera las biyecciones siguientes operando en la *recta proyectiva* o el *plano analítico* (ej. 61)

$$x \mapsto x, \qquad x \mapsto 1/x, \qquad x \to 1-x, \qquad x \mapsto 1/(1-x),$$
$$x \to (x-1)/x, \qquad x \mapsto x/(x-1).$$

Demostrar que este conjunto es un grupo para la composición de las aplicaciones. Comparar este grupo al grupo de las isometrías del triángulo (ej. 60) y al grupo $\mathfrak{S}_3$ de las permutaciones de un conjunto de tres elementos. Determinar todos sus subgrupos.

63. Buscar todos los grupos de n elementos para $n = 5$ o 6 (utilizar el mismo método que en el ejercicio del § 71). Ver igualmente el ejercicio 81.

64. Demostrar que un grupo en el que para todo x, $x^2 = e$, es conmutativo.

65. Demostrar que los axiomas de un grupo son equivalentes a los axiomas siguientes: G_1) la ley es asociativa; G_2') existe un elemento neutro por la izquierda (resp. por la derecha); G_3') todo elemento tiene un inverso por la izquierda (resp. por la derecha) (V. ej. 52, capítulo 3).

66.* Demostrar que todo conjunto E provisto de una ley asociativa tal que para todo a de E las traslaciones por la izquierda y por la derecha γ_a y δ_a sean suprayectivas es un grupo (V. ej. 47, capítulo 3).

67*. A y B son dos subgrupos de un grupo G, se considera el subgrupo S engendrado por A ∪ B.

a) Demostrar que todo elemento de S se obtiene como producto de una sucesión finita de elementos pertenecientes alternativamente a A y B.

b) Demostrar que $S = AB$ si y sólo si $AB = BA$. ¿Qué sucede si A o B es un subgrupo invariante?

c) Deducir de lo anterior que el conjunto de los subgrupos de G ordenado por inclusión es un retículo (capítulo 1, ej. 24). La misma pregunta para el conjunto de los subgrupos invariantes de G (precisar entonces el sup y el inf para cada pareja de subgrupos invariantes).

68. Sea G un grupo no conmutativo de centro C y G′, el conjunto de los automorfismos inferiores f_a de G ($f_a(x) = axa^{-1}$).

a) Demostrar que G′ es un grupo para la composición de las aplicaciones.

b) Demostrar que $a \mapsto f_a$ define un homomorfismo de G sobre G′.

c) Demostrar que G′ es isomorfo a G/C (ver § 77, ej. 4, y § 76, ej. 2).

69. Sea A una parte no vacía de un grupo G, se llama *normalizador* de A en G el conjunto N(A) descrito por los x de G tales que $xA = Ax$, y *centralizador* de A en G, el conjunto C(A) descrito por los x de G tales que para todo a de A, $xa = ax$.
Demostrar que N(A) es un subgrupo de G y C(A) es subgrupo invariante de N(A). ¿Cuál es el centralizador de G?

70. Sea f un homomorfismo de un grupo G *sobre* un grupo G′ y H′ un subgrupo invariante de G′.
Demostrar que $H = f^{-1}(H')$ es un subgrupo invariante de G y que φ (resp. φ') al ser el homomorfismo canónico de G (resp. G′) sobre G/H (resp. G′/H′), existe un isomorfismo i de G/H sobre G′/H′ tal que $\varphi' \circ f = i \circ \varphi$.

b) Se supone que K es un subgrupo invariante de G tal que $f(K) \subset H'$, siendo ψ el homomorfismo canónico de G sobre G/K y ψ' de G′ sobre G′/H′, demostrar que existe un homomorfismo de h de G/K sobre G′/H′ tal que $\varphi' \circ f = h \circ \psi$.

71*. Se considera un grupo G de orden $2n$.

a) Demostrar que todo subgrupo H de G de orden n es invariante.

b) Si existe en G dos subgrupos de orden n, H_1 y H_2 tales que $H_1 \cap H_2 = \{e\}$, n es igual a 2 y G tiene una estructura determinada, dar su tabla. ¿Cuántos subgrupos de orden 2 posee G?

72*. a) Sea G_1 y G_2 dos grupos, de elementos neutros respectivos e_1 y e_2 y $G'_1 = G_1 \times \{e_2\}$; se sabe (§ 80, *b*) que G'_1 es un subgrupo invariante del grupo $G_1 \times G_2$; demostrar que $G_1 \times G_2$ es isomorfo al grupo producto $G'_1 \times [(G_1 \times G_2)/G'_1]$.

b) Dado un grupo G y H un subgrupo invariante de G demostrar que para que G sea isomorfo a H × (G/H), hace falta que G/H sea isomorfo a un grupo invariante de G: ¿sucede así para el grupo aditivo **Z** y uno de sus subgrupos propios?

c) Con las mismas notaciones que en **b**) demostrar que G es isomorfo a H × (G/H) si y sólo si existe un homomorfismo de G sobre H cuya restricción a H sea la identidad.

73. Se considera el grupo $K = \{e, a, b, c\}$ definido por $a^2 = b^2 = c^2 = e$, $bc = a$, $ca = b$, $ab = c$ (grupo de *Klein*).

a) Designando, respectivamente, por Sx, Sy, Sz las simetrías con relación a los tres ejes de un triedro trirrectángulo, demostrar que el grupo engendrado por esas simetrías es isomorfo a K.

b) Demostrar que K es isomorfo al grupo aditivo $(\mathbf{Z}/2\mathbf{Z}) \times (\mathbf{Z}/2\mathbf{Z})$.

c) Encontrar todos los endomorfismos y automorfismos de K. Demostrar que el grupo de los automorfismos de K es isomorfo a $\mathfrak{S}_3$.

74. Consideremos el grupo aditivo $G = \mathbf{Z}/n\mathbf{Z}$.

a) Demostrar que un endomorfismo f de G está determinado por el valor $f(1)$. ¿Cuántos endomorfismos de G existen?

b) Determinar todos los automorfismos de G. Demostrar que hay $\varphi(n)$, donde $\varphi(n)$ es el número de los enteros p primos con n tales que $1 \leq p \leq n$.

75. Tenemos dos grupos G_1 y G_2 de elementos neutros respectivos e_1 y e_2, se considera un endomorfismo φ_1 de G_1 y un endomorfismo de φ_2 de G_2.

a) Demostrar que la aplicación f del grupo $G_1 \times G_2$ en sí mismo definida por $f(x_1, x_2) = [\varphi_1(x_1), \varphi_2(x_2)]$ es un endomorfismo de $G_1 \times G_2$.

b) Sea g un endomorfismo del grupo $G_1 \times G_2$; ¿en qué condición g es un endomorfismo del tipo estudiado en la pregunta a)? (se introducirá $G'_1 = G_2 \times \{e_2\}$ y $G'_2 = \{e_1\} \times G_2$ y se estudiará la estabilidad eventual de G'_1 y G'_2 por g). Estudiar el caso en que $G_1 = G_2 = \mathbf{Z}/2\mathbf{Z}$ (ver ej. 73).

76. Sea G el grupo engendrado por a y b tales que $a^4 = e$, $b^2 = e$, $ab = ba^3$; demostrar que $a^2b = ba^2$ y $a^3b = ba$; deducir de ello que G está descrito por los ocho elementos $e, a, a^2, a^3, ab, ba, a^2b, b$. Formar su tabla. Determinar los subgrupos de G.

77*. Sea G el grupo de las isometrías de un polígono regular de n vértices de centro O (V. ej. 60). Se designa por la rotación de centro O y de ángulo $2\pi/n$ y por s la simetría con relación al radio que une el centro con un vértice determinado A_0. (Se escribirá fg en lugar de $f \circ g$, ver § 82, ej. 2).

a) Demostrar que (e identidad) $r^n = s^2 = e$, $rsr = s$.

b) Construir las tablas en el caso $n = 2$ (el polígono regular es un segmento ! se encuentra de nuevo el grupo de *Klein*, ej. 73); $n = 3$ (ver ej. 62); $n = 4$ (ver ej. 76).

c) En el caso general se demostrará que para todo entero racional z, $r^z s = sr^{-z}$, después que todo elemento del grupo puede escribirse en la forma $r^x \cdot r^y$ (x e y enteros naturales tales que $0 \leq x \leq n - 1$, $0 \leq y \leq 1$). De lo anterior deducir que r y s engendran G y que G es de orden $2n$; se le llama grupo *diedral* y se le designa por $\mathfrak{D}_{2n}$.

d) Demostrar que $\mathfrak{D}_{2n}$ es isomorfo al grupo que describe el conjunto

$$(\mathbf{Z}/n\mathbf{Z}) \times (\mathbf{Z}/2\mathbf{Z})$$

provisto de una ley de composición interna que se determinará.

78. Se considera el grupo G engendrado por a y b verificando $a^4 = b^4 = e$, $a^2 = b^2$, $aba = b$ ($a \neq b$) (llamado grupo *cuaterniónico*).

a) Demostrar que este grupo G es de orden 8, formar su tabla. Demostrar que no es isomorfo a $\mathfrak{D}_4$ (ej. 77).

b) Demostrar que todo subgrupo de G es distinguido y que la intersección de dos subgrupos distinguidos de G es distinto de $\{e\}$.

79. p es un número primo, se designa por M el conjunto $\mathbf{Z}/p\mathbf{Z} - \{\dot{0}\}$ provisto de la multiplicación de los enteros módulos p y por A al grupo aditivo $\mathbf{Z}/(p-1)\mathbf{Z}$.

a) Demostrar que M es un grupo abeliano.

b) Se supone $p = 11$; demostrar que 2 engendra M, luego que M es un grupo cíclico de 10 elementos.
Determinar todos los enteros módulos 11 que engendran M.

c) Demostrar, siempre para $p = 11$, que A y M son isomorfos. (El resultado es general, ver capítulo 11, ej. 351, **c**).)

80. Sea G un grupo conmutativo.

a) Si a es de orden p y b de orden q, p y q primos entre sí, demostrar que $c = ab$ es de orden pq (siendo A el subgrupo engendrado por a y B el subgrupo engendrado por b, se mostrará que $A \cap B = \{e\}$).

b) Si a_i es de orden p_i $(1 \leq i \leq m)$, los m enteros naturales, siendo p_i primos entre sí dos a dos, demostrar que $a_1, a_2, \ldots, a_m$, es de orden $p_1, p_2, \ldots, p_m$.

81. Buscar nuevamente los grupos G de orden n $(2 \leq n \leq 7)$ utilizando el hecho de que el orden de un elemento y el orden de un subgrupo son divisores de n (§§ 74 y 83); observar también que para n primo, G es cíclico (§ 83, ej. 6).

82*. Sea G un grupo conmutativo finito de orden p^n, p es un número primo.

a) Demostrar que todo elemento es de orden p^m $(0 < m \leq n)$ y que hay elementos de orden p (utilizar el teorema final del § 74, **b**; seguidamente siendo a de orden $p^m = r$, $m > 1$, considerar el elemento $a^{r/p}$).

b) Demostrar que todo subgrupo es de orden p^m $(0 < m \leq n)$ y que para todo m existe al menos un subgrupo de orden p^m (El resultado es trivial para $n = 1$, razonar por recurrencia considerando el subgrupo A engendrado por un elemento a de orden p y el grupo cociente G/A que es de orden p^{n-1} si G es de orden p^n (§ 76, ej. 5); sea, hipótesis de recurrencia, H′ un subgrupo de orden $m-1$ $(0 < m-1 \leq n-1)$ de G/A, considerar la imagen recíproca H de H′ en el homomorfismo canónico $G \to G/A$.)

83*. Sea G un grupo conmutativo finito de un número finito de elementos. Sea

$$m = p_1^{r_1} \ldots p_k^{r_k}$$

la descomposición de m en factores primos. Se pone en todo lo que sigue

$$q_i = p_i^{r_i}, \qquad m_i = m/q_i \qquad (i = 1, \ldots, k)$$

y se designa por G_i el conjunto de los elementos x de G tales que $x^{q_i} = e$, donde e designa el elemento neutro de G.

a) Demostrar que G_i es un subgrupo de G.

b) Demostrar que existen enteros $u_1, \ldots, u_k$ tales que $u_1m_1 + \ldots + u_km_k = 1$.

c) Para todo x de G se pone $x_i = x^{u_im_i}$ $(i = 1, \ldots, k)$, en donde u_i nos los da la pregunta **b**). Demostrar que x_i pertenece a G_i y que $x = x_1, \ldots, x_k$.

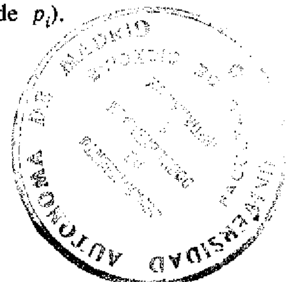
d) Se considera la aplicación del grupo producto $G_1 \times \dots \times G_k$ en G definida por $(x_1, \dots, x_k) \to x_1, \dots, x_k$; demostrar que esta aplicación es un isomorfismo de grupos.

e) Demostrar que G_i es de orden q_i (para $i = 1, \dots, k$).

f) Desarrollar los resultados precedentes en el caso en que G es el grupo aditivo $\mathbf{Z}/m\mathbf{Z}$ de los enteros módulo m.

(Examen MGP, París 1960).

(para **b**) ver § 100, teorema 5'; para **c**) utilizar el hecho que para todo x de G, $x^m = e$; para **d**) se demostrará primero que la aplicación es un homomorfismo suprayectivo, se demostrará seguidamente que es inyectivo buscando su núcleo; para **e**) se demostrará que el orden de x_i es una potencia de p_i, se considerará seguidamente el grupo G_i/X_i, X_i engendrado por un elemento x_i de G_i —V. ej. 80 **b**)— y se demostrará por recurrencia que el orden de G_i es una potencia de p_i).



84. Se considera el diagrama conmutativo siguiente (§ 15)

$$\begin{array}{ccc} E & \xrightarrow{f} & E \\ \varphi\downarrow & & \downarrow\varphi \\ E & \xrightarrow[f']{} & E \end{array}$$

en donde f y φ son dos biyecciones de E; demostrar que $f \mapsto f'$ es un automorfismo interior de $\mathfrak{S}_E$. Se dice que la biyección f' es *transmutada* de f por φ: f y f' son *conjugadas* en el grupo $\mathfrak{S}_E$ (§ 89, ej. 2).

a) En el grupo de desplazamientos del plano, ¿cuál es la transformada de f por un desplazamiento φ en los casos siguientes: 1. f es una traslación; 2. f es una rotación?

b) En el grupo de las isometrías del plano, ¿cuál es la transformada de una simetría axial de eje D por una rotación cuyo centro está sobre D?

c) En el grupo, que opera en el plano analítico **C** (ej. 61), engendrado por las inversiones, demostrar que la transformación de una inversión por una inversión es también una inversión.

85. Descomponer la permutación $p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 7 & 8 & 9 & 4 & 5 & 2 & 1 & 6 \end{pmatrix}$ en producto de ciclos.

¿Cuál es el orden de p en $\mathfrak{S}_9$? calcular p^{1000}.

86. Para $n > 3$ se designa por $\mathfrak{A}'_n$ el subgrupo de $\mathfrak{S}_n$ engendrado por los $n-2$ ciclos (1 2 3), (1 2 4), ..., (1 2 n).

a) Demostrar que $\mathfrak{A}'_n$ es un subgrupo de $\mathfrak{S}_n$.

b) Demostrar que (1 2) (i j) e (i j) (1 2) (i y j distintos) pertenecen a $\mathfrak{A}'_n$.

c) Demostrar que $\mathfrak{A}'_n = \mathfrak{A}_n$ (observar que toda permutación de $\mathfrak{A}_n$ puede escribirse

$$p = t_1\, t_2 \cdots t_{2p-1}\, t_{2p} = t_1\, t_0\, t_0\, t_2\, t_0 \cdots t_0\, t_{2p-1}\, t_0\, t_0\, t_{2p}$$

con $t_0 = (1, 2)$ y utilizar **b**).

87*. Se dice que un grupo G es *simple* si no posee otro subgrupo invariante que $\{e\}$ y G. Se propone demostrar que para $n > 4$ el grupo alternado $\mathfrak{A}_n$ es simple. Sea H un subgrupo invariante de $\mathfrak{A}_n$ distinto de $\{e\}$.

a) Demostrar que todos los ciclos de orden 3 son conjugados en $\mathfrak{S}_n$ (§ 89, ej. **2**). Deducir de lo anterior que si H contiene un ciclo de orden 3, $H = \mathfrak{A}_n$.

b) Demostrar que si H contiene un ciclo de orden estrictamente superior a 3 contiene también un ciclo de orden 3; sea

$$p = (a_1, a_2, a_3, a_4, \ldots, a_m) \in H, \qquad q = (a_1, a_2, a_3) \in \mathfrak{S}_n$$

p^{-1} y $q^{-1}pq$ pertenecen a H, luego también $p^{-1}q^{-1}pq$; pero esta última permutación es un ciclo de orden 3.

c) Demostrar, en fin, que es imposible que H sólo contenga productos (en número par) de transposiciones.

d) *Deducir de* **a)**, **b)**, **c)** que $H = \mathfrak{A}_n$.

88. Tenemos un grupo abeliano G representado aditivamente, se dice que una relación de orden es compatible con la adición en G si para todo x de G

$$a \prec b \Rightarrow a + x \prec b + x,$$

se dice entonces que G es un *grupo ordenado.*

a) Demostrar que toda relación de orden compatible con la adición en G es de la forma $b - a \in P$, P es el conjunto de los elementos superiores al elemento neutro **0**. Demostrar que

$$(1) \qquad P + P \subset P, \qquad P \cap (-P) = \{0\}.$$

b) Recíprocamente, demostrar que si P es una parte de G verificando las dos relaciones (1), la relación $b - a \in P$ es una relación de orden compatible con la adición en G.

Demostrar que el orden es total si y sólo si $G = P \cup (-P)$.

c) En el grupo aditivo $\mathbf{R}^2$, demostrar que se puede tomar como parte P los puntos *de un ángulo saliente cuyo vértice es* O (comprendidos sus lados). ¿A qué parte P de $\mathbf{R}^2$ corresponden los órdenes definidos sobre $\mathbf{R}^2$ en el § 24?

d) Tomemos de nuevo todo este ejercicio con notación multiplicativa (G siempre abeliano). Demostrar que en el grupo $\mathbf{Q}^*_+$ se puede tomar para P verificando $1 \in P$, $PP \subset P$ y $P \cap (P^{-1}) = \{1\}$ el conjunto $\mathbf{N}^*$; explicar la relación $ba^{-1} \in \mathbf{N}^*$; ¿el orden obtenido es total?

89. Se dice que un grupo abeliano ordenado (representado aditivamente) es *arquimediano* si, cualesquiera que sean $a > 0$ y $b > 0$, *existe* n *de* **N** *tales que* $na > b$. **Z** *es uno* de tales grupos (§ 60).

Demostrar que el grupo aditivo $\mathbf{R}^2$ totalmente ordenado por $(a, a') \leq (b, b')$ si y solamente si $a < b$ o ($a = b$ y $a' \leq b'$), no es arquimediano (si $a' > 0$ y $b > 0$ para todo n: $n(0, a') < (b, b')$).

5

ANILLOS Y CUERPOS

I. Anillos. Primeras propiedades

90. Estructura de anillo. Ejemplos

DEFINICIÓN. — *Un conjunto* A *provisto de una adición y de una multiplicación posee una* estructura de anillo *para estas operaciones si:*

— A posee una estructura de grupo conmutativo para la adición.

— La multiplicación es asociativa.

— La multiplicación es distributiva por la izquierda y por la derecha con respecto a la suma.

Se dice también que A es un *anillo* para la adición y la multiplicación consideradas, o simplemente que A es un anillo si no da lugar a confusión.

Los axiomas de la estructura de anillo son, pues, distinguiendo los axiomas relativos a cada operación, y los axiomas de "compatibilidad" entre las dos operaciones (ver § 68),

$$\left\{\begin{array}{lll} A_1 & (\forall a, b, c \in A) & (a+b)+c = a+(b+c) \\ A_2 & (\exists \varepsilon \in A)(\forall a \in A) & a+\varepsilon = \varepsilon + a = a \\ A_3 & (\forall a \in A)(\exists a' \in A) & a + a' = a' + a = \varepsilon \qquad a' = -a \\ A_4 & (\forall a, b \in A) & a+b = b+a \end{array}\right.$$

$$A_5 \quad (\forall a, b, c \in A) \qquad (ab)c = a(bc)$$

$$\left\{\begin{array}{lll} A_6 & (\forall a, b, c \in A) & a(b+c) = ab+ac \\ A_7 & (\forall a, b, c \in A) & (b+c)a = ba+ca. \end{array}\right.$$

b) Si la multiplicación es conmutativa, se dice que el anillo es *abeliano* o *conmutativo*.

En un anillo no necesariamente conmutativo se dice que dos elementos particulares a, b son *permutables* si $ab = ba$.

Se llama *elemento central* del anillo A, un elemento que permuta con todo elemento del anillo; el conjunto de los elementos centrales es el *centro* del anillo.

Si la multiplicación tiene un elemento neutro e se dice que el anillo es *unitario*.

ε (elemento neutro de la adición) se llama el *elemento cero* o también el cero del anillo, e (elemento neutro de la multiplicación, si existe) se llama *elemento unidad*, o la unidad del anillo.

Si no es fácil que surja alguna confusión, se les puede representar, respectivamente, por 0 y 1; pero en ciertos casos la notación 1 en lugar de e puede ser incorrecta (ver § 97).

Siendo A un *anillo unitario*, si dado a de A existe a' de A tal que $aa' = a'a = e$, se dice que a es *inversible en* A, o que a es un elemento *inversible*(16) del anillo.

EJEMPLOS Y EJERCICIOS

1. **Z** y **Z**/n**Z** son anillos conmutativos unitarios. ¿Cuáles son los elementos inversibles de **Z**? Demostrar que en el segundo anillo $\widehat{n-1}$ es inversible.

2. El conjunto de los polinomios en x (por ejemplo, de coeficientes reales), provisto de la adición y de la multiplicación de los polinomios, es un anillo conmutativo unitario.

3. Demostrar que un anillo contiene al menos un elemento ε y que existe un solo anillo con único elemento. Se le llama el *anillo cero* (se tiene $\varepsilon + \varepsilon = \varepsilon$, $\varepsilon^2 = \varepsilon$) o anillo nulo.

4. El conjunto $\mathcal{F}(\mathbf{R}, \mathbf{R})$ de las aplicaciones de **R** en **R** provisto de las leyes

$$(f, g) \to s = f + g \qquad (\forall x \in \mathbf{R}) \qquad s(x) = f(x) + g(x)$$
$$(f, g) \to p = fg \qquad (\forall x \in \mathbf{R}) \qquad p(x) = f(x)g(x)$$

es un anillo conmutativo unitario; determinar el elemento cero y el elemento unidad de este anillo. (No confundir $p = fg$ y $c = f \circ g$.)

5. De una manera más general, demostrar que el conjunto $\mathcal{F}(E, A)$ de las aplicaciones de un conjunto cualquiera E en un anillo A, provisto de las dos operaciones $f + g$ y fg definidas como hemos hecho anteriormente (ej. 4), es un anillo. (Generalización de las consideraciones del § 54.) En particular $\mathcal{F}(A, A)$ tiene una estructura de anillo.

91. Reglas de cálculo sobre un anillo. Anillo de integridad. Anillo unitario

a) La multiplicación es distributiva en relación a la resta; en efecto, cualesquiera que sean a, b, c

$$a(c - b) + ab = a[(c - b) + b] = ac$$
$$(c - b)a + ba = [(c - b) + b]a = ca$$

(16) Algunos autores llaman «unidades» a estos elementos inversibles, nosotros no lo haremos para evitar confusiones con «el elemento unidad».

de donde

(1) $$a(c - b) = ac - ab$$
(1') $$(c - b)a = ca - ba.$$

En (1) y (1') $b = c$ da para todo a

(2) $$a\varepsilon = \varepsilon a = \varepsilon.$$

Por otro lado, como en todo grupo representado aditivamente (ver § 66 y tabla del § 70) se puede definir na para todo n de **Z** y todo a de A, en particular $0a = \varepsilon$, que comparada con la fórmula (2) se ve que *en general no hay ningún inconveniente en utilizar el símbolo* 0 *a la vez para el cero de* **Z** *y para el cero de* A, *lo que haremos en adelante.* Pondremos $A^* = A - \{0\}$.

En (1) y (1') $c = 0$ da cualesquiera que sean a y b

$$a(-b) = -(ab) \qquad (-b)a = -(ba)$$

de donde

$$(-a)(-b) = -a(-b) = -[-(ab)] = ab$$

y para n estrictamente positivo

$$n \text{ par} \qquad (-a)^n = a^n$$
$$n \text{ impar} \qquad (-a)^n = -a^n.$$

b) La *recíproca* de propiedad traducida por (2) $a0 = 0a = 0$, es decir, "$ab = 0$ *implica* $a = 0$ o $b = 0$", es verdadera o falsa según el anillo considerado: es verdadera en **Z** (§ 62), es falsa en **Z/6Z** (§ 18, ej. 3, y § 53, ej. 1), pues $\dot{2} \cdot \dot{3} = \dot{3} \cdot \dot{4} = \dot{0}$. De lo anterior se deducen las siguientes definiciones:

DEFINICIÓN. — *Cuando en un anillo existen elementos a, b tales que*

$$a \neq 0 \qquad b \neq 0 \qquad ab = 0$$

se dice que a y b son verdaderos divisores de cero o simplemente divisores de cero.

Se llama anillo de integridad *o* anillo íntegro *un anillo conmutativo, no reducido a cero y desprovisto de divisores de cero.*

c) Se llama elemento *nilpotente* todo elemento a tal que existe un entero natural no nulo verificando

$$a^n = 0$$

se ve que todo a nilpotente no nulo es un divisor de cero.

d) Como lo hemos visto anteriormente, para todo n de **Z** y todo a de A se puede definir na, pero en general n no pertenece a A. Sin embargo, si el anillo A es *unitario*, se puede escribir este elemento na bajo la forma de un producto de dos elementos de A. Sea e el elemento unidad, las reglas de

cálculo en un anillo y la convención $0 = \varepsilon$ muestran que

$$
\begin{array}{ll}
n > 0 & na = a + a + \ldots + a = ea + ea + \ldots + ea = (ne)a \\
n = 0 & 0a = 0 = (0e)a \\
n < 0 & na = (-n)(-a) = (-n)[e(-a)] = (-n)[(-e)a] = [(-n)(-e)]a = (ne)a
\end{array}
$$

se mostraría igualmente $na = a(ne)$, luego para todo n de **Z** y todo a de **A**

$$na = (ne)a = a(ne)$$

na es, pues, el producto de dos elementos ne y a del anillo (que son permutables).

e) Consideremos las dos sumas $A = \sum_{i \in I} a_i$, $B = \sum_{j \in J} b_j$ con $I = [1, p]$, $J = [1, q]$ y su producto AB. Las reglas de cálculos anteriores nos permiten escribir

$$
\begin{aligned}
AB &= (a_1 + \ldots + a_i + \ldots + a_p)(b_1 + \ldots + b_j + \ldots + b_q) \\
&= (a_1b_1 + \ldots + a_1b_j + \ldots + a_1b_q) + \ldots + (a_ib_1 + \ldots + a_ib_j + \ldots + a_ib_q) \\
&\quad + \ldots + (a_pb_1 + \ldots + a_pb_j + \ldots + a_pb_q) = \sum_{i \in I}\Big(\sum_{j \in J} a_ib_j\Big) \\
&= (a_1b_1 + \ldots + a_ib_1 + \ldots + a_pb_1) + \ldots + (a_1b_j + \ldots + a_ib_j + \ldots + a_pb_j) \\
&\quad + \ldots + (a_1b_q + \ldots + a_ib_q + \ldots + a_pb_q) = \sum_{j \in J}\Big(\sum_{i \in I} a_ib_j\Big) \\
&= a_1b_1 + \ldots + a_ib_j + \ldots + a_pb_q = \sum_{(i,j) \in I \times J} a_ib_j
\end{aligned}
$$

luego

$$AB = \Big(\sum_{i \in I} a_i\Big)\Big(\sum_{j \in J} b_j\Big) = \sum_{i \in I}\Big(\sum_{j \in J} a_ib_j\Big) = \sum_{j \in J}\Big(\sum_{i \in I} a_ib_j\Big) = \sum_{(i,j) \in I \times J} a_ib_j$$

se puede, pues, escribir suprimiendo los paréntesis

$$AB = \sum_{i \in I}\sum_{j \in J} a_ib_j = \sum_{j \in J}\sum_{i \in I} a_ib_j$$

el resultado está puesto bajo la forma de un "sigma doble". Observemos que si $I = J$ los índices de las a y las b describen I *independientemente* uno del otro y deben, pues, representarse por *letras distintas*. Así la notación

$$AB = \Big(\sum_{i \in I} a_i\Big)\Big(\sum_{i \in I} b_i\Big)$$

es *correcta*, pero *insegura*, pues

$$AB = \sum_{(i,j) \in I \times I} a_ib_j \neq \sum_{i \in I} a_ib_i = \sum_{(i,j) \in \Delta} a_ib_j$$

en donde Δ representa la diagonal de $I \times I$.

Todo esto puede extenderse al producto de n sumas finitas

$$A_1 = \sum_{i_1 \in I_1} a^1_{i_1} \dots A_k = \sum_{i_k \in I_k} a^k_{i_k} \dots A_n = \sum_{i_n \in I_n} a^n_{i_n}$$

k, que es un *índice* y no un exponente, indica el número de la suma que pertenece $a^k_{i_k}$, i_k describe el conjunto finito I_k. Tendremos

$$P = A_1 \dots A_k \dots A_n = \sum_{i_1 \in I_1} \dots \sum_{i_k \in I_k} \dots \sum_{i_n \in I_n} a^1_{i_1} \dots a^k_{i_k} \dots a^n_{i_n}$$

$$= \sum_{(i_1, \dots, i_k, \dots, i_n) \in I_1 \times \dots \times I_k \times \dots \times I_n} a^1_{i_1} \dots a^k_{i_k} \dots a^n_{i_n}$$

hemos descrito un "sigma *n-étuple*". Naturalmente si $I_1 = \dots = I_n = I$, la observación anterior es aún válida.

EJEMPLOS Y EJERCICIOS

1. Demostrar que los divisores de cero de un anillo A no son regulares para la multiplicación. Los elementos inversibles de un anillo no son jamás divisores de cero.
2. ¿En qué condición el anillo $\mathbf{Z}/n\mathbf{Z}$ es íntegro?
3. Demostrar que el anillo definido en el ejercicio 4, párrafo precedente, no es íntegro.
4. Demostrar que si A, diferente del anillo cero (§ 90, ej. 3) es unitario, el elemento unidad e es diferente de cero.

92. Anillo conmutativo. Fórmula del binomio

Las reglas de cálculo clásicas en **Z**, **Q** o **R** no son siempre válidas en un anillo cualquiera; por ejemplo, si, en un anillo no conmutativo, a y b no son permutables, se tiene

$$(a + b)^2 = (a + b)(a + b) = a^2 + ab + ba + b^2 \neq a^2 + 2ab + b^2$$
$$(a + b)(a - b) = a^2 - ab + ba - b^2 \neq a^2 - b^2.$$

a) Si el anillo es conmutativo las fórmulas clásicas relativas a $(a + b)^2$, $(a + b)^3$, ..., $(a + b)(a - b)$ son aún válidas. De una manera más general calcularemos $(a + b)^n$ en un anillo conmutativo para n entero positivo

$$(a + b)^n = (a + b)(a + b) \dots (a + b) \qquad (n \text{ factores})$$
$$= \gamma^0_n a^n + \gamma^1_n a^{n-1} b + \dots + \gamma^m_n a^{n-m} b^m + \dots + \gamma^n_n b^n$$

γ^0_n, γ^1_n, ..., γ^m_n, ..., γ^n_n que son *enteros naturales* no nulos y verifican visiblemente

$$\gamma^0_n = \gamma^n_n = 1 \qquad (\text{hay un solo término } a^n \text{ y un solo término } b^n)$$
$$\gamma^m_n = \gamma^{m-n}_n \qquad ((a + b)^n = (b + a)^n).$$

Por otra parte, la relación (para $n \geq 1$)

$$(a + b)^n = (a + b)^{n-1}(a + b)$$

permite que escribamos para $1 \leq m \leq n$

$$\gamma_n^m = \gamma_{n-1}^{m-1} + \gamma_{n-1}^m.$$

Luego si formamos el cuadro de los coeficientes γ_n^m las leyes de formación son las mismas que las del cuadro de los C_n^m (ver § 41, *b*), luego

$$\gamma_n^m = C_n^m = \binom{n}{m} = \frac{n!}{m!(n-m)!} = \frac{n(n-1)\dots(n-m+1)}{m!}$$

$$\boxed{\begin{aligned}(a + b)^n &= C_n^0 a^n + C_n^1 a^{n-1}b + \dots + C_n^m a^{n-m}b^m + \dots + C_n^n b^n \\ &= \sum_{m=0}^{n} C_n^m a^{n-m} b^m\end{aligned}}$$

esta fórmula se llama "*fórmula del binomio*", de donde el nombre de *coeficientes binomiales* dados a los enteros naturales C_n^m.

OBSERVACION

Naturalmente, la fórmula del binomio es aún válida en un anillo no conmutativo si a y b son *permutables* y más generalmente en un conjunto E provisto de una adición (asociativa y conmutativa) y de una multiplicación asociativa, para la cual a y b son permutables, y distributiva con relación a la suma.

b) Se puede generalizar la fórmula del binomio. En todo anillo conmutativo se demuestra que (ver ej. 1)

$$(a_1 + a_2 + \dots + a_m)^n = \sum \frac{n!}{p_1!p_2!\dots p_m!}(a_1)^{p_1}\dots(a_m)^{p_m}$$

la suma del segundo miembro se extiende a todas las sucesiones $p_1, p_2, \dots, p_m$, de enteros naturales tales que $\sum_{i=1}^{m} p_i = n$.

En particular,

$$(a_1 + a_2 + \dots + a_m)^2 = \sum_{1\leq i\leq m} (a_i)^2 + 2 \sum_{1\leq i<j\leq m} a_i a_j$$

$$(a_1 + a_2 + a_3 + \dots + a_m)^3 = \sum_{1\leq i\leq m} (a_i)^3 + 3\sum_{i\neq j}(a_i)^2 a_j + 6\sum a_i a_j a_k$$

en el último $\sum$, i, j, k son tres elementos de $[1, m]$ dos a dos distintos.

EJERCICIOS

1. Demostrar la fórmula que da la potencia n-ésima de una suma utilizando el ejercicio 36 (fin del capítulo 2).

2. Dada una progresión aritmética

$$a_1,\ a_2 = a_1 + r,\ \dots\ a_i = a_{i-1} + r,\ \dots\ a_n = a_{n-1} + r$$

se pone

$$S_p = (a_1)^p + (a_2)^p + \dots + (a_n)^p.$$

Desarrollar para $2 \leq i \leq n$ $(a_i + r)^{p+1}$ mediante la fórmula del binomio. Sumando las igualdades obtenidas, encontrar una relación entre $S_0 = n$, S_1, ..., S_p.

Aplicar los cálculos precedentes al caso $a_1 = r = 1$ y demostrar que

$$S_1 = 1 + 2 + \dots + n = \frac{n(n+1)}{2}$$

$$S_2 = 1^2 + 2^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

$$S_3 = 1^3 + 2^3 + \dots + n^3 = \frac{n^2(n+1)^2}{4} = (S_1)^2.$$

3. Calcular $\displaystyle\sum_{k=1}^{k=1.000} k(k+1)(k+4)$.

4. Determinar la fórmula (ver ej. 6, § 41) $\displaystyle\sum_{m=0}^{m=n} C_n^m = 2^n$ desarrollando $(1+1)^n$. Se designa por P la suma de los coeficientes binomiales en los que m es par y por I la suma de estos coeficientes en los que m es impar. Demostrar que $P = I = 2^{n-1}$.

5. De la relación $(a+b)^{p+q} = (a+b)^p(a+b)^q$ deducir una relación entre los coeficientes binomiales igualando en los dos miembros los coeficientes de $a^n b^{p+q-n}$ $(0 \leq n \leq p+q)$.

6. Calcular

$$(C_n^0)^2 + \dots + (C_n^m)^2 + \dots + (C_n^n)^2$$
$$C_n^1 + 2C_n^2 + \dots + mC_n^m + \dots + nC_n^n.$$
$$C_n^0 + (1/2)C_n^1 + \dots + (1/m+1)C_n^m + \dots + (1/n+1)C_n^n.$$

93. Subanillos

DEFINICIÓN. — *Se llama subanillo de un anillo* A *toda parte no vacía* B *de* A *estable para las leyes de* A *y tal que la estructura inducida sobre* B *por estas leyes sea una estructura de anillo.*

Un subanillo B de A es, pues, un *subgrupo del grupo aditivo* A y una parte estable de A para la multiplicación. Recíprocamente una parte B de A verificando estas dos condiciones es un subanillo de A. En efecto, la multipli-

cación inducida sobre la parte estable B de A es asociativa y distributiva po la izquierda y por la derecha con relación a la suma. De donde, utilizando e teorema del § 72, *c*:

TEOREMA. — *Para que una parte no vacía* B *de un anillo* A *sea un subanillo de* A *es necesario y suficiente que*

$$(a \in B \quad y \quad b \in B) \Rightarrow (a - b \in B \quad y \quad ab \in B).$$

Si A es conmutativo o íntegro, lo es también el subanillo B; pero A pued ser unitario sin que B lo sea, como lo muestra el resultado siguiente:

TEOREMA. — *Los subanillos del anillo* **Z** *son los conjuntos* $n\mathbf{Z}$ *(n entero racion cualquiera).*

En efecto, los subgrupos del grupo **Z** son los conjuntos $n\mathbf{Z}$, que son visibl mente estables para la multiplicación; constatamos que para $n \geq 2$ estos sub anillos no son unitarios, aunque **Z** lo sea.

Finalmente, como para los subgrupos (ver § 73), se demostrará fácilment que toda intersección de subanillos de A es un subanillo de A: se podr pues, definir el menor subanillo conteniendo una parte no vacía X de A es la intersección de todos los subanillos de A conteniendo X, se le llama e *subanillo engendrado por* X.

EJERCICIOS

1. Determinar los subanillos de $\mathbf{Z}/6\mathbf{Z}$, $\mathbf{Z}/n\mathbf{Z}$.
2. Demostrar que el conjunto de las aplicaciones reales de variable real nulas par x_0 describen un subanillo de $\mathfrak{F}(\mathbf{R}, \mathbf{R})$ definido en el ejercicio 4 del § 90.
3. Demostrar que el centro C de un anillo A es un subanillo de A.
4. Demostrar que el conjunto de los elementos de un anillo A que permutan co un elemento *a* de A —o con cada elemento de una parte X de A— es un subanillo de A
5. Demostrar que **Z** es un subanillo de **Q** considerado como anillo.

II. Ideales. Homomorfismos de anillos

94. Noción de ideal. Ejemplos

a) Una relación de equivalencia R será compatible con la estructura d un anillo A si y sólo si para todo x de A es

$$(1) \qquad a \equiv b \pmod{R} \Rightarrow a + x \equiv b + x \pmod{R}$$
$$(2) \qquad a \equiv b \pmod{R} \Rightarrow xa \equiv xb \pmod{R}$$
$$(2') \qquad a \equiv b \pmod{R} \Rightarrow ax \equiv bx \pmod{R}.$$

Los resultados del § 75 demuestran que una relación R que verifique (1 es de la forma $a - b \in B$, siendo B un subgrupo del grupo aditivo A. Pa

una tal relación las condiciones (2) y (2′) se escriben $x(a-b)\in B$ y $(a-b)x\in B$; ahora bien, si B es un subgrupo cualquiera no existe razón alguna para que $x(a-b)$ y $(a-b)x$ pertenezcan a B para todo x de A. Nos vemos así conducidos a distinguir subgrupos particulares del grupo aditivo A.

DEFINICIÓN. — *Se llama* ideal por la izquierda (*resp.* por la derecha) *de un anillo* A, *todo subgrupo* I_g (*resp.* I_d) *del grupo aditivo* A, *estable para la multiplicación por la izquierda* (*resp. por la derecha*) *por un elemento cualquiera del anillo.*

Todo ideal que es a la vez ideal por la derecha e ideal por la izquierda de A *se llama* ideal bilateral *de* A.

Es decir,

$$(\forall a,\ b\in I_g)\quad a-b\in I_g;\qquad (\forall a,\ b\in I_d)\quad a-b\in I_d$$
$$(\forall a\in I_g)(\forall x\in A)\quad xa\in I_g;\qquad (\forall a\in I_d)(\forall x\in A)\quad ax\in I_d.$$

Naturalmente, en un anillo conmutativo A, estas tres nociones: ideal por la izquierda, ideal por la derecha e ideal bilateral coinciden, se dice que I es un *ideal* del anillo conmutativo A. Observamos que un ideal de A es un subanillo de A.

EJEMPLOS

1. Siendo A un anillo cualquiera $\{0\}$ y A son ideales bilaterales de A; un ideal distinto de $\{0\}$ y A se llama *ideal propio de* A.

2. Busquemos los ideales del anillo **Z**, hay que buscarlos entre los subgrupos del grupo aditivo **Z**, es decir, los conjuntos $n\mathbf{Z}$. Pues cualquiera que sean n, a y x de **Z**

$$na\in n\mathbf{Z}\Rightarrow x(na)=n(xa)\in n\mathbf{Z}.$$

Los ideales de **Z** *son los conjuntos* $n\mathbf{Z}$.

3. Se verificará fácilmente que, siendo a un elemento fijo de A, aA es un ideal por la derecha de A y Aa un ideal por la izquierda de A.

b) Como para los subgrupos (§ 73) y los subanillos (§ 93), toda intersección de ideales por la izquierda de un anillo A es un ideal por la izquierda de A: se podrá definir el menor ideal por la izquierda conteniendo una parte no vacía X de A: es la intersección de todos los ideales por la izquierda de A conteniendo X; se le llama el ideal por la izquierda de A *engendrado* por X. Se definiría igualmente el ideal por la derecha engendrado por X y el ideal bilateral engendrado por X.

Busquemos, por ejemplo, el ideal por la izquierda engendrado por un elemento fijo a de A. Como subgrupo del grupo aditivo de A debe contener na para todo entero racional n, como ideal por la izquierda de A debe contener xa para todo x de A, luego todo elemento de la forma $xa+na$, x y n describiendo, respectivamente, A y **Z**. Ahora bien, estos elementos describen un ideal por la izquierda de A. En efecto,

$$(xa+na)-(x'a+n'a)=(x-x')a+(n-n')a$$
$$y(xa+na)=(yx)a+(yn)a=(yx+yn)a,$$

$x-x'$, $yx+yn$ son elementos de A y $n-n'$ de **Z**.

Si A *posee un elemento unidad* $na = (ne)a$ (§ 91, *d*) y ne es un elemento de A, así como $x + ne$; luego:

Para todo anillo unitario el ideal por la izquierda engendrado por a está descrito por los elementos xa, cuando x describe A, *es* Aa. (Generalmente este resultado no es verdadero en un anillo desprovisto de elemento unidad, ver ej. 4.)

Si, además, el anillo es conmutativo A$a = a$A el ideal engendrado por a se representa (a), se dice que es *principal*. Por ejemplo, $A = (e)$.

DEFINICIÓN. — *Se llama anillo principal, todo anillo íntegro unitario y en él todo ideal es principal.*

Así, **Z** es un *anillo principal*.

De una manera más general, se verificará fácilmente que para un *anillo conmutativo unitario* el ideal engendrado por $a_1, a_2, \ldots, a_n$ está descrito por los elementos

$$u_1a_1 + u_2a_2 + \ldots + u_na_n$$

de donde $u_1, u_2, \ldots, u_n$ describen A. Este ideal se representa $(a_1, a_2, \ldots, a_n)$ (no confundir esta notación con la de un n-étuple, elemento de A^n).

EJERCICIOS

1. Entre los subanillos estudiados en los ejercicios del 1 al 5 del § 93, determinar los que son ideales.

2. Demostrar que el conjunto de ideales por la izquierda, por ejemplo, de un anillo A es un retículo (capítulo 1, ej. 24), inf $(I_1, I_2) = I_1 \cap I_2$, sup (I_1, I_2) *es el ideal por la izquierda engendrado por* $I_1 \cup I_2$, es $I_1 + I_2$. El mismo resultado para los ideales por la derecha y los ideales bilaterales.

3. Siendo Y una parte no vacía de un anillo A el conjunto de los x de A tales que $xy = 0$ para todo y de Y es un ideal por la izquierda de A llamado *anulador por la izquierda* de Y.

Definir igualmente el anulador por la derecha de Y. Si A no tiene divisores de cero y si Y contiene al menos un elemento no nulo, demostrar que existe un solo anulador.

4. En $A = 2\mathbf{Z}$, demostrar que el ideal I engendrado por 4 no es igual al ideal J descrito por $4x$, si x describe A. Demostrar que las inclusiones siguientes son estrictas: $IA \subset I$, $JA \subset J$.

OBSERVACION

Sea B un subanillo de A, es evidente que todo ideal de A, contenido en B, es un ideal de B. Pero la recíproca es falsa en general: por ejemplo, $n\mathbf{Z}$ $(n \neq 0)$ es un ideal de **Z**, y no es un ideal de **Q** considerado como anillo.

95. Relaciones de equivalencia compatibles con una estructura de anillo. Anillo cociente

Las consideraciones del párrafo precedente, que han conducido a la definición de los ideales, muestran que las únicas relaciones de equivalencia de

finidas sobre un anillo A, compatibles con la adición y compatibles por la izquierda (resp. por la derecha) con la multiplicación, son de la forma $a - b \in I_g$ (resp. $a - b \in I_d$), I_g (e I_d) es un ideal por la izquierda (por la derecha) de A.

Luego, dada una relación de equivalencia R definida sobre A, si queremos definir una adición y una multiplicación en A/R por las relaciones

$$(\dot{x}) + (\dot{y}) = \left(\dot{\widehat{x + y}}\right) \qquad (\dot{x})(\dot{y}) = \left(\dot{\widehat{xy}}\right)$$

es necesario y suficiente (ver § 53) que R sea compatible con la adición y la multiplicación de A, es decir, que R sea de la forma

$$a - b \in I$$

donde I es un *ideal bilateral* de A.

Se escribe esta relación

$$a \equiv b \qquad (\text{mod I})$$

que se lee "*a congruente con b módulo el ideal bilateral* I".

Las relaciones precedentes proporcionan al conjunto A/R una estructura de anillo; en efecto, A/R es un grupo aditivo (ver § 75); por otra parte, la multiplicación es asociativa

$$[(\dot{x})(\dot{y})](\dot{z}) = (\dot{\widehat{xy}})(\dot{z}) = \dot{\widehat{xyz}} = (\dot{x})(\dot{\widehat{yz}}) = (\dot{x})[(\dot{y})(\dot{z})]$$

y distributiva por la izquierda y por la derecha con respecto a la multiplicación; por ejemplo,

$$(\dot{x})[(\dot{y}) + (\dot{z})] = (\dot{x})\left(\dot{\widehat{y + z}}\right) = \dot{\widehat{x(y + z)}} = \left(\dot{\widehat{xy + xz}}\right) = (\dot{\widehat{xy}}) + (\dot{\widehat{xz}})$$
$$= (\dot{x})(\dot{y}) + (\dot{x})(\dot{z})$$

de donde:

TEOREMA Y DEFINICIÓN. — *Las únicas relaciones de equivalencia* R *compatibles con la estructura de un anillo* A *son de la forma* $a - b \in I$, *siendo* I *un ideal bilateral de* A.

El conjunto cociente de A *por* R *tiene una estructura de anillo, se le llama el* anillo cociente *de* A *por* I *y se le representa* A/I.

Así, en **Z** la relación $a \equiv b$ (módulo n) se escribe también $a - b \in n\mathbf{Z}$, es compatible con la adición y la multiplicación en **Z**, como lo hemos visto de manera elemental en el § 64; el anillo cociente que hemos utilizado varias veces en los ejemplos y en los ejercicios se representa $\mathbf{Z}/n\mathbf{Z}$ o también algunas veces $\mathbf{Z}/(n)$.

96. Homomorfismos, isomorfismos de anillos

a) Las consideraciones de los §§ 56, 57 y 69 nos permiten enunciar:

DEFINICIÓN. — *Una aplicación f de un anillo* A *en un anillo* A′ *es un* homomorfismo *de anillos si y sólo si*

$$(\forall\ x,\ y \in \mathrm{A}) \qquad f(x+y) = f(x) + f(y) \qquad f(xy) = f(x)f(y).$$

Un homomorfismo de A *en sí mismo se llama un* endomorfismo *del anillo* A. *Si f es biyectiva, se dice que f es un* isomorfismo *de* A *sobre* A′. *Un isomorfismo de* A *sobre sí mismo se llama* automorfismo *del anillo* A.

Por ejemplo, si I es un ideal bilateral de A la aplicación de A sobre A/I definida por

$$x \rightarrow f(x) = \dot{x}$$

es por definición del anillo A/I un homomorfismo de anillos (suprayectivo), se le llama *homomorfismo canónico* de A sobre A/I.

b) TEOREMA 1. — *La composición de dos homomorfismos (resp. isomorfismos) de anillos es un homomorfismo (resp. isomorfismo) de anillos.*

TEOREMA 2. — *Si f es un homomorfismo de un anillo* A *en un anillo* A′:

1. $f(0) = 0, \ f(-x) = -f(x)$.
2. $f(\mathrm{A})$ *es un subanillo de* A′.
3. $\mathrm{N} = f^{-1}(0)$ *es un ideal bilateral de* A, *se le llama el núcleo del homomorfismo.*

1. Se ha demostrado en el § 56, teorema 2, y recordado en el § 77, teorema 2.

2. Sabemos ya que $f(\mathrm{A})$ es un subgrupo del grupo aditivo A′ (§ 77, teorema 2). Mostremos que, cualesquiera que sean x' e y' de $f(\mathrm{A})$, $x'y'$ pertenecen a $f(\mathrm{A})$: sea x e y dos elementos de A tales que $f(x) = x'$, $f(y) = y'$, tendremos $x'y' = f(x)f(y) = f(xy)$; luego, según el teorema del § 93 sobre los subanillos, $f(\mathrm{A})$ es un subanillo de A′.

3. Sabemos ya que $f^{-1}(0)$ es un subgrupo del grupo aditivo A (§ 77, teorema 2). $f(a) = 0$ implica cualquiera que sea x de A: $f(ax) = f(a)f(x) = 0$, $f(x) = 0$, igualmente $f(xa) = 0$; luego $f^{-1}(0)$ es un ideal bilateral de A.

COROLARIO. — *El homomorfismo de anillos f*: $\mathrm{A} \rightarrow \mathrm{A}'$ *es inyectivo si y sólo si* $f^{-1}(0) = \{0\}$.

OBSERVACIONES

1. Es inútil suponer que A′ es un anillo, es suficiente suponer que A′ es un conjunto provisto de una adición y de una multiplicación, el teorema 2 del § 56 nos permite afirmar que $f(\mathrm{A})$ es una parte estable de A′, se demostrará fácilmente que para la adición y la multiplicación inducidas por las de A′, tiene una estructura de anillo.

2. Ciertas propiedades algebraicas de A son válidas también para $f(\mathrm{A})$ (ver § 56, t. 2), se demostrará que si A es conmutativo $f(\mathrm{A})$ también lo es, que si A tiene un

elemento unidad e, $e' = f(e)$ es elemento unidad de $f(A)$. En fin, que si x es inversible en A, también lo es $f(x)$ en $f(A)$ y $[f(x)]^{-1} = f(x^{-1})$.

Por el contrario, el homomorfismo canónico de $\mathbf{Z}$ sobre $\mathbf{Z}/n\mathbf{Z}$ muestra que A puede ser íntegro sin que $f(A)$ lo sea.

TEOREMA 3. — *Si f es un isomorfismo del anillo* A *sobre el anillo* A′, f^{-1} *es un isomorfismo de* A′ *sobre* A, *se dice entonces que* A *y* A′ *son anillos isomorfos.*

c) Siendo f un homomorfismo de un anillo A en un anillo A′, consideremos la descomposición canónica de f (ver § 19, *b*) asociada a la relación de equivalencia R sobre A

$$f(x) = f(y) \Leftrightarrow f(x - y) = 0 \Leftrightarrow x - y \in f^{-1}(0) = N.$$

Acabamos de ver (teorema 2) que N es un ideal bilateral de A; tenemos, pues, la descomposición

$$A \xrightarrow{s} A/N \xrightarrow{b} f(A) \xrightarrow{i} A'$$
$$f = i \circ b \circ s$$

s es el *homomorfismo canónico* del anillo A sobre el anillo cociente A/N; i, *la inyección canónica* $(x' \to x')$, es un homomorfismo de anillos; respecto a la *biyección b* se ve fácilmente cómo para un grupo (ver § 78) que es un isomorfismo. En efecto, dadas dos clases $\dot{x}$ e $\dot{y}$, módulo N, de representantes respectivos x e y en A, tenemos por definición

$$b(\dot{x}) = f(x) \qquad b(\dot{y}) = f(y),$$

de donde

$$b(\dot{x} + \dot{y}) = b\left(\dot{\widehat{x + y}}\right) = f(x + y) = f(x) + f(y) = b(\dot{x}) + b(\dot{y})$$
$$b(\dot{x}\dot{y}) \quad = b\,(\dot{\widehat{xy}}) \quad = f(xy) \quad = f(x)f(y) \quad = b(\dot{x})b(\dot{y}).$$

EJERCICIOS

1. Siendo a un elemento inversible fijado de un anillo unitario A, demostrar que la aplicación de A en A definida por $x \to axa^{-1}$ es un automorfismo del anillo A.

2. Demostrar que $\mathbf{Z}[\sqrt{2}]$ (ver capítulo 3, ej. 59) tiene una estructura de anillo para la adición y la multiplicación. Sea A el conjunto $\mathbf{Z} \times \mathbf{Z}$ provisto de las dos operaciones

$$(a, a') + (b, b') = (a + b, a' + b')$$
$$(a, a')(b, b') = (ab + 2bb', ab' + a'b)$$

demostrar que la aplicación de $\mathbf{Z}[\sqrt{2}]$ en A definida por

$$a + a'\sqrt{2} \to (a, a')$$

es un isomorfismo. Deducir de ello que A es un anillo conmutativo unitario.

97. Característica de un anillo

Sea A un anillo con elemento unidad e, consideremos el subgrupo monogéneo A', del grupo aditivo de A engendrado por e y la aplicación f de **Z** en A' definida por $f(n) = ne$; las reglas de cálculo en un grupo aditivo muestran que, cualesquiera que sean los enteros racionales m y n,

$$f(m + n) = (m + n)e = me + ne = f(m) + f(n)$$
$$f(mn) = (mn)e = m(ne) = (me)(ne) = f(m)f(n)$$

luego f, suprayectiva por definición, es un homomorfismo suprayectivo del anillo A sobre $A' = f(A)$ que es un subanillo de A.

Los resultados obtenidos en el § 83 sobre los grupos monógenos (en notación multiplicativa, pero son válidos con notación aditiva) y el hecho de que f es un homomorfismo de anillos nos permite enunciar el:

TEOREMA Y DEFINICIÓN.—*Sea* A *un anillo, con un elemento unidad* e, *si la aplicación* f *de* **Z** *en* A *definida por* $f(n) = ne$ *es inyectiva, el único entero* p *tal que* $pe = 0$ *es* $p = 0$ *y* $f(A)$ *es un anillo isomorfo a* **Z**, *se dice que* A *es de característica nula.*

Si la aplicación f *no es inyectiva, existe un menor entero* $p > 0$, *tal que* $pe = 0$ *y* $f(A)$ *es un anillo isomorfo a* $\mathbf{Z}/p\mathbf{Z}$, *en este caso se dice que* A *es de característica* $p > 0$.

Si A es de característica nula, no hay ningún inconveniente en *identificar* n con **Z** y ne con A, es decir, en representar el *elemento unidad* de A por el mismo símbolo que el elemento unidad 1 de **Z**: lo que permite decir que *todo anillo unitario de característica nula contiene* **Z**.

Por el contrario, cuando la característica p de A es no nula esta identificación puede conducir a resultados absurdos (por ejemplo, el anillo finito $\mathbf{Z}/p\mathbf{Z}$ contiene **Z**).

Observemos, en fin, que en un anillo A de característica $p > 0$, cualquiera que sea a de A: $pa = (pe)a = 0a = 0$.

EJERCICIOS

1. Sea p un número primo, demostrar que para todo entero q tal que $1 \leq q \leq p - 1$, C_p^q es divisible por p.

Deducir que en un anillo de características p primo $(a + b)^p = a^p + b^p$, luego que $(a - b)^p = a^p - b^p$ (poner $a = a - b + b$) y finalmente que

$$(a_1 + a_2 + \ldots + a_k)^p = (a_1)^p + \ldots + (a_k)^p.$$

2. Siendo p un entero natural primo y k un entero natural cualquiera, demostrar que $k^p \equiv k \pmod p$. (Se calculará $(\dot{1} + \dot{1} + \ldots + \dot{1})^p$, k términos, en $\mathbf{Z}/p\mathbf{Z}$; ver otra demostración de este teorema debida a FERMAT, § 104, ej. 3).

III. Divisibilidad en un anillo. Estudio particular de Z

En esta sección vamos a considerar de nuevo el estudio de la divisibilidad en el anillo **Z**, ya estudiado en Matemáticas Elementales; lo haremos con la noción de *ideal*, lo que nos conducirá a nociones válidas en anillos más generales A, que supondremos, a menudo, *unitarios y conmutativos*. En un tal anillo el ideal engendrado por a es $a\mathrm{A} = \mathrm{A}a$, es principal; se le representa (a) (§ 94, *c*). Sólo demostraremos los teoremas fundamentales, enunciando los secundarios y los corolarios.

En un último párrafo indicaremos por qué esta teoría de la divisibilidad es relativamente sencilla en **Z** y daremos algunas indicaciones sobre la complejidad que puede presentar en anillos más generales.

98. Divisibilidad e inclusión de los ideales principales. Aplicaciones

a) En un anillo conmutativo unitario A tenemos

$$b \mid a \Leftrightarrow [(\exists\, q \in \mathrm{A}) \quad a = bq] \Leftrightarrow (a) \subset (b)$$

luego $(a) = (b)$ implica la existencia de q y q' tales que $b = aq$, $a = bq'$, es decir, $a = aqq'$, luego si $a \neq 0$ y si, además, A es íntegro, $qq' = e$: q es un elemento inversible de A.

Consideremos de una manera general el conjunto U de los *elementos inversibles* de un anillo A unitario, U es estable por la multiplicación, pues si u y v son inversibles uv también lo es y $(uv)^{-1} = v^{-1}u^{-1}$ (§ 48, *b*), la ley inducida sobre U por la multiplicación de A es asociativa; en fin, U contiene e y conteniendo u contiene u^{-1}, de donde:

TEOREMA 1. — *El conjunto* U *de los elementos inversibles de un anillo unitario* A *es estable para la multiplicación; es un grupo para la ley inducida sobre* U *por la multiplicación de* A.

EJEMPLOS Y EJERCICIOS

1. En **Z** el grupo U es el grupo multiplicativo $\{1, -1\}$ (§ 65).

2. En el anillo de los polinomios en x de coeficientes reales, el grupo U es el grupo multiplicativo $\mathbf{R}^*$ (§ 90, ej. 2).

3. Determinar el grupo U para los anillos siguientes

$$\mathcal{F}(\mathbf{R}, \mathbf{R}) \ (\text{§ } 90, \text{ ej. } 4), \qquad \mathbf{Z}/5\mathbf{Z}, \qquad \mathbf{Z}/12\mathbf{Z}, \qquad \mathbf{Z}/n\mathbf{Z} \ (\text{§ } 95).$$

Las consideraciones del principio del párrafo nos conducen al resultado siguiente:

TEOREMA 2. — *En un anillo conmutativo unitario*

$$(a) \subset (b) \Leftrightarrow b \mid a.$$

En un anillo íntegro unitario

$$(a) = (b) \Leftrightarrow b = au$$

donde u es un elemento de A; *se dice entonces que a y b son elementos* asociados.

OBSERVACIONES

1. Este teorema es particularmente válido para todo anillo principal (§ 94, *b*), luego en **Z**.

2. Los elementos asociados a a en **Z** son, pues, $\pm a$. Luego **Z** es un anillo principal, se puede siempre hacer corresponder a un ideal $I \neq \{0\}$ de **Z** un entero único $a > 0$ tal que $I = (a)$.

3. En **Z** si b divide $a \neq 0$

$$0 < |b| \leq |a| \Leftrightarrow (a) \subset (b)$$

se tendrá cuidado con este «*cambio*» del sentido de los órdenes: así, $2 < 6$, pero 2 tiene «muchos» más múltiplos que 6, luego $(2) \supset (6)$ (estrictamente).

b) Sea a un entero de **Z** divisible solamente por ± 1 y $\pm a$, observemos que al admitir 0 todo entero por divisor (§ 63) a es no nulo, luego $(a) \neq \{0\}$:

Si $a = \pm 1$, $(a) = \mathbf{Z}$.

Si $a \neq \pm 1$, la desigualdad $(a) \subset \mathbf{Z}$ es estricta; sea I un ideal de **Z** conteniendo estrictamente (a), según las observaciones anteriores y el teorema 2, existe b único tal que $I = (b)$ y $0 < |b| < |a|$, b al dividir a, según la propiedad de a, $b = 1$ e $I = \mathbf{Z}$.

Este estudio nos conduce a enunciar la definición general siguiente:

DEFINICIÓN 1. — *Se dice que un ideal* I *de un anillo conmutativo es* maximal *si* I *es un elemento maximal (para la inclusión) del conjunto de los ideales de* A, *distintos de* A.

Luego si I es un ideal maximal de A

$$I \neq A \quad \text{e} \quad (I \subset I' \Rightarrow I' = I \quad \text{o} \quad I' = A).$$

Por otra parte, todo elemento a de un anillo íntegro, unitario, es divisible por e, a y los elementos que le son asociados, lo que nos conduce a distinguir los elementos que no tienen otros divisores añadiendo una condición suplementaria (veremos por qué).

DEFINICIÓN 2. — *Un elemento p de un anillo íntegro unitario* A *es* extremal *si es distinto de un elemento inversible y si sólo es divisible por los elementos inversibles o los elementos que le son asociados. En* **Z** *un elemento extremal se llama* número primo.

Luego en **Z** si p es primo hemos demostrado que (p) es maximal; recíprocamente sea (p) un ideal maximal, desde luego $(p) \neq \mathbf{Z}$, luego $p \neq \pm 1$; en consecuencia, no existe ningún divisor q de p tal que $1 < |q| < |p|$ si no $(q) \neq \mathbf{Z}$ contendría estrictamente (p), de donde:

TEOREMA 3. — *En* **Z** *el ideal* (p) *es maximal si y sólo si p es primo*

OBSERVACIONES

1. El término «elemento *extremal*» se reemplaza a menudo por el término «elemento *primo*» (como en **Z**) o por el término «elemento *irreducible*» (por ejemplo, en los anillos de polinomios; ver capítulo 11).

2. Observemos que la condición p primo es no inversible, es decir, $|p| \neq 1$ en **Z** es esencial para la validez del teorema 3 (también lo será para los anillos principales, donde este teorema es también cierto; ver ej. 105).

3. Puesto que todo elemento a divide a 0 ($0a = 0$), *un elemento extremal es siempre no nulo.*

c) DEFINICIÓN 3. — *Dos elementos a y b de un anillo unitario* A *son* extraños *o* primos entre sí, *si sólo tienen como divisores comunes elementos inversibles de* A. *Se dice también que a (resp. b) es extraño a b (resp. a).*

Observemos que al dividir todo elemento a cero ($0a = 0$) dos elementos extraños son no nulos.

En **Z** dos enteros extraños sólo tienen como divisores comunes 1 y -1. Se preferirá el término "*extraño*" al término "*primos entre sí*" para no confundirlos con el término "*entero primo*".

DEFINICIÓN 3′. — *Los elementos $a_1, a_2, \ldots, a_n$ de una familia finita de elementos de un anillo unitario* A *son* extraños en su conjunto *o* primos entre sí en su conjunto *si sólo tienen como divisores comunes elementos inversibles de* A.

OBSERVACION

No se debe confundir una familia de elementos *extraños en su conjunto* y una familia de elementos *extraños dos a dos*: esta segunda familia es un caso particular de la primera.

99. Máximo común divisor de dos elementos de Z

a) Si a y b son dos enteros racionales *no nulos*, consideremos el ideal $I = (a, b)$, engendrado por a y b; está descrito por los elementos de la forma $ax + by$, en donde x e y describen **Z**; es, pues, el ideal $(a) + (b)$. Como en **Z** todo ideal es principal, existe $D > 0$ tal que

$$(a, b) = (D)$$

(en efecto, D no puede ser nulo sin que (a, b) sea el ideal cero). Perteneciendo D a I, existen u y v tales que

$$au + bv = D \tag{1}$$

Igualmente al pertenecer a y b a I, existe a' y b' tales que

$$a = a'D \qquad b = b'D \tag{2}$$

según (1) todo divisor común a a y b divide D, según (2) todo divisor de D divide a y b; existe, pues, *un único máximo común divisor estrictamente posi-*

tivo de a y de b, o sea, D; se dirá simplemente que D es el m. c. d. de a y de b y se le representará $D(a, b)$.

TEOREMA 4. — *Dados dos enteros racionales no nulos a y b, existe un único máximo común divisor estrictamente positivo D de a y de b y*

$$(a) + (b) = (D).$$

Además, existen dos enteros racionales u y v tales que

$$au + bv = D. \tag{1}$$

OBSERVACION

u y v no son únicos; en efecto, cualquiera que sea el entero k, (1) implica

$$a(u + kb) + b(v - ka) = D.$$

La igualdad (1) anterior y la definición 3 del § 98 permiten enunciar los teoremas 5 y 6:

TEOREMA 5. — *Dados dos enteros racionales a y b, las propiedades siguientes son equivalentes:*

1. *a y b son extraños.*
2. *El m. c. d. de a y b es 1.*
3. *Existen enteros racionales u y v tales que (igualdad de Bezout)*

$$au + bv = 1. \tag{3}$$

4. *Para todo entero racional z, existen enteros racionales x e y tales que $ax + by = z$.*

Naturalmente, las parejas (u, v) y (x, y) no son únicas (ver ej. 1).
Si p es primo, para todo a de $\mathbf{Z}$, $D(a, p) = 1$ o $|p|$, de donde:

COROLARIO. — *Todo entero primo no extraño con el entero a divide a.*

TEOREMA 6. — *Si $D(a, b)$ es el m. c. d. de dos enteros racionales no nulos tenemos:*

1. *Para todo c no nulo*

$$D(ac, bc) = |c|\, D(a, b).$$

2. *Para todo divisor común d de a y b*

$$D(ad^{-1}, bd^{-1}) = |d^{-1}|\, D(a, b).$$

3. *Siendo d un divisor común de a y de b, para que $|d|$ sea su m. c. d. es necesario y suficiente que ad^{-1} y bd^{-1} sean extraños.*

b) TEOREMA 7. — *Las tres proposiciones siguientes son equivalentes:*

1. *a y b son extraños.*
2. *Para todo x, a divide bx implica a divide x.*
3. *Para todo y, b divide ay implica b divide y.*

(1) es simétrico; basta, pues, demostrar que (1) implica (2) (teorema de GAUSS). En efecto, $D(a, b) = 1$ implica $D(ax, bx) = |x|$, a divide ax y bx, luego a su m. c. d. que es $|x|$.

Demostremos que (2) implica (3): supongamos que b divide ay, $ay = bq$, según (2) a que divide bq divide q, luego $q = aq'$, $ay = baq'$, luego $y = bq'$. Se demostrará igualmente que (3) implica (2).

Demostremos en fin que (2), por ejemplo, implica (1); sea d un divisor común de a y b

$$(a = a'd \quad \text{y} \quad b = b'd) \Rightarrow ab' = ba'$$

según (2) a dividiendo ba' divide a', pero a' divide a, luego $a' = \pm a$ y $d = \pm 1$, lo que nos dice que a y b son extraños.

COROLARIOS:

1. *Si a es extraño con b y con c, lo es también con bc y con b^n $(n > 0)$.*
2. *Dadas dos familias finitas (a_i) y (b_j) de enteros racionales, si cada a_i es extraño con cada b_j, el producto de los a_i es extraño con el producto de los b_j.*
3. *Si a es divisible por cada elemento de una familia finita (b_i) de enteros racionales extraños dos a dos, es divisible por el producto de los b_i.*
4. $D(a^n, b^n) = [D(a, b)]^n$, $(n > 0)$.
5. *Si p primo divide ab y no divide a, divide b.*

Este corolario 5 nos conduce a una noción importante. Sea p un entero tal que para *todo* producto ab, p no pueda dividir ab sin dividir a o b. Sea a un divisor de p, se tiene $p = aa'$, dividiendo p a aa' debe dividir a o a':

— Si p divide a, como a divide p: $a = \pm p$.

— Si p no divide a, entonces divide a'; se tiene, pues, $a' = pq$, de donde $aa' = p = pqa$, es decir, si $p \neq 0$, $aq = 1$, resulta $a = q = \pm 1$.

Luego si p es no nulo, los únicos divisores de p son ± 1, $\pm p$. Interpretemos la condición que verifica p con ayuda del ideal (p); la condición $y \in (x)$ equivale a x divide y; por lo tanto,

$$[ab \in (p) \quad \text{y} \quad a \notin (p)] \Rightarrow b \in (p).$$

Lo que nos conduce a la definición general y al teorema siguiente:

DEFINICIÓN 4. — *Dado un anillo conmutativo* A, *se dice que el ideal* I *es primo si y sólo si*

$$[ab \in \mathrm{I} \quad \text{y} \quad a \notin \mathrm{I}] \Rightarrow b \in \mathrm{I}.$$

TEOREMA 8. — *Un entero p no nulo y distinto de ± 1 es primo si y sólo si el ideal (p) es primo.*

COROLARIOS:

1. *Si p primo divide $a_1, a_2, \ldots, a_n$, divide al menos a uno de los factores.*
2. *Si p primo divide a^n $(n > 0)$, divide a.*

EJERCICIOS

1. *a*) Demostrar que si (u, v) es una pareja correspondiente a la fórmula (3) (igualdad de BEZOUT) todas las otras son $(u + kb,\ v - ka)$, siendo k un entero cualquiera.

b) Si a y b son dos enteros estrictamente positivos extraños, demostrar que si $a > 2$ y $b > 2$, existe una pareja única (u, v) verificando la fórmula 3 y tal que $|u| < b/2$, $|v| < a/2$. Estudiar el caso en que a o b sea igual a 2.

c) Si a y b son dos enteros extraños, demostrar que todo entero z tal que $|z| < |ab|$ puede ponerse en la forma $z = ax + by$ con $|x| < |b|$ e $|y| < |a|$ y esto de una o dos maneras. Dar un ejemplo en que haya dos expresiones, como las anteriores, de z.

2. *Algoritmo de Euclides* para determinar el m. c. d. de a y b.

a) Al efectuar las divisiones euclídeas (se supondrá $b > 0$)

$$\begin{array}{ll} a = bq_0 + r_0 & 0 \leq r_0 < b \\ b = r_0 q_1 + r_1 & 0 \leq r_1 < r_0 \\ r = r_1 q_2 + r_2 & 0 \leq r_2 < r_1 \\ r_{n-1} = r_n q_{n+1} + r_{n+1} & 0 \leq r_{n+1} < r_n \end{array}$$

se llegará forzosamente a un primer resto $r_{n-1} = 0$, demostrar que

$$D(a, b) = D(q, r) \ldots = D(r_{n-1}, r_n) = r_n.$$

b) Demostrar por recurrencia la existencia de enteros u_k, v_k tales que $r_k = au_k + bv_k$, deducir una nueva demostración de la fórmula (1) $D = au + bv$ y un modo de cálculo de una pareja (u, v).

3. Si a, b, c son tres enteros dados (a y b extraños) resolver en números enteros la ecuación $ax + by = c$.

4. Demostrar que todo ideal I de un anillo conmutativo **A** es *primo* si y sólo si el anillo A/I es *íntegro.* ¿El ideal cero de A puede ser un ideal primo de A?

100. Máximo común divisor de una familia finita de elementos de Z

Si $a_1, a_2, \ldots, a_n$ son enteros racionales no todos nulos, el ideal $\mathbf{I} = (a_1, a_2, \ldots, a_n)$ engendrado por $a_1, a_2, \ldots, a_n$ está descrito por los elementos de la forma

$$x_1 a_1 + x_2 a_2 + \ldots + x_n a_n,$$

donde $x_1, x_2, \ldots, x_n$ describen **Z**; es, pues, el ideal $(a_1) + (a_2) + \ldots + (a_n)$.

Como en **Z** todo ideal es principal, existe $D > 0$ tal que

$$(a_1, a_2, \ldots, a_n) = (D)$$

(D no puede ser nulo sin que I sea el ideal cero, lo que es imposible, si uno al menos de los a_i no es nulo).

Luego D pertenece a I; existen, pues, enteros $u_1, u_2, \ldots, u_n$ tales que

$$(1') \qquad a_1 u_1 + a_2 u_2 + \ldots + a_n u_n = D$$

igualmente a_i pertenece a I; hay, pues, un entero a'_i tal que

$$(2') \qquad a_i = a'_i D \qquad (1 \leq i \leq n)$$

según (1′) todo divisor común a $a_1 \ldots a_n$ es divisor de D, según (2′) todo divisor de D divide cada a_i, luego existe *un divisor común máximo estrictamente positivo único* de $a_1, a_2, \ldots, a_n$, o sea, D; se dirá simplemente que D es el m. c. d. de $a_1, a_2, \ldots, a_n$ y se le representará $D(a_1, a_2, \ldots, a_n)$.

TEOREMA 4'. — *Dados los elementos* $a_1, a_2, \ldots, a_n$ *de una familia finita de enteros racionales no todos nulos, hay un máximo común divisor estrictamente positivo único* D *y*

$$(a_1) + (a_2) + \ldots + (a_n) = (D).$$

Además, existen enteros racionales $u_1, u_2, \ldots, u_n$ *tales que*

$$(1') \qquad a_1u_1 + a_2u_2 + \ldots + a_nu_n = D.$$

La igualdad (1') anterior y la definición (3') del § 98 permiten enunciar los teoremas (5') y (6'):

TEOREMA 5'. — *Dada una familia finita de enteros racionales* $a_1, a_2, \ldots, a_n$, *las propiedades siguientes son equivalentes:*

1. $a_1, a_2, \ldots, a_n$ *son extraños en su conjunto.*
2. *El m. c. d. de* $a_1, a_2, \ldots, a_n$ *es* 1.
3. *Existen enteros racionales* $u_1, u_2, \ldots, u_n$ *tales que (igualdad de Bezout)*

$$(3') \qquad a_1u_1 + a_2u_2 + \ldots + a_nu_n = 1.$$

4. *Para todo entero racional* y, *existen enteros racionales* $x_1, x_2, \ldots, x_n$ *tales que* $y = a_1x_1 + \ldots + a_nx_n$.

TEOREMA 6'. — *Si* $D(a_1, a_2, \ldots, a_n)$ *es el m. c. d. de una familia de enteros racionales no todos nulos, entonces:*

1. *Para todo* $b \neq 0$

$$D(a_1b, a_2b, \ldots, a_nb) = |b|\, D(a_1, a_2, \ldots, a_n).$$

2. *Para todo divisor común* d *de* $a_1, a_2, \ldots, a_n$

$$D(a_1d^{-1}, \ldots, a_nd^{-1}) = |d^{-1}|\, D(a_1, a_2, \ldots, a_n).$$

3. *Para que un divisor común* d *de* $a_1, a_2, \ldots, a_n$ *sea tal que* $|d|$ *sea el m. c. d. de* $a_1, a_2, \ldots, a_n$ *es necesario y suficiente que* $a_1d^{-1}, \ldots, a_nd^{-1}$ *sean extraños en su conjunto.*

EJERCICIO

Demostrar que en **Z**, $(a, b) \to D(a, b)$ es una ley interna asociativa, deducir de esta ley un método para encontrar el m. c. d. de una familia finita de enteros racionales.

101. Mínimo común múltiplo de dos o varios elementos de Z

Si a y b son dos enteros racionales no nulos, el conjunto de sus múltiplos comunes es el conjunto de los elementos comunes al ideal (a) y al ideal (b); es, pues, la intersección de (a) y de (b); siendo este ideal principal existe, por tanto, un entero $M > 0$ único tal que

$$(a) \cap (b) = (M)$$

(en efecto, M no puede ser nulo, ya que $ab \neq 0$ pertenece a la intersección), todo múltiplo común a a y b es, en consecuencia, múltiplo de M, de donde:

TEOREMA 9. — *Dados dos enteros racionales no nulos a y b, existe un mínimo común múltiplo estrictamente positivo* M *único y*

$$(a) \cap (b) = (M).$$

M *se llama el* mínimo común múltiplo *de a y b (m. c. m.), se le designa* M(a, b).

Igualmente para los elementos de una familia finita de enteros racionales no nulos, tendremos:

TEOREMA 9'. — *Dados los elementos $a_1, a_2, \ldots, a_n$ de una familia de enteros racionales no nulos, existe un mínimo común múltiplo estrictamente positivo* M *único y*

$$(a_1) \cap (a_2) \ldots \cap (a_n) = (M).$$

M *se llama el* mínimo común múltiplo *de $a_1 \ldots a_n$ (m. c. m.), y se le representa mediante* M$(a_1, a_2, \ldots, a_n)$.

EJERCICIOS

1. Si a y b son dos enteros racionales no nulos, demostrar que $|ab| = D(a, b)M(a, b)$ (poner $a = Da'$, $b = Db'$, $m = aa'' = bb''$ para todo múltiplo común a a y b).
2. Demostrar que en **Z**, $(a, b) \to M(a, b)$ es una ley interna asociativa, deducir un método para encontrar el m. c. m. de una familia finita de enteros racionales.

102. Descomposición de un entero racional en factores primos

a) Si a es primo admite al menos un divisor primo: el mismo; *supongamos a no primo,* admite, pues, al menos un divisor b distinto de ± 1 y $\pm a$. Si b divide a, $\pm b$ divide $\pm a$, se puede, pues, suponer a y b positivos; tenemos entonces

$$a = bq \quad \text{y} \quad 1 < b < a.$$

Existe un número finito de estos divisores $b > 1$, el menor entre ellos es evidentemente primo, luego:

TEOREMA 10. — *Todo entero racional admite un divisor primo. Dicho de otra manera, todo ideal de* **Z** *está contenido en un ideal maximal de* **Z**.

b) Sea a un entero racional no nulo, admite un divisor primo $p_1 > 0$ (teorema 10): $a = p_1a_1$, si a_1 no es primo admite un divisor primo $p_2 > 0$, luego $a_1 = p_2a_2$, $a = p_1p_2a_2$. Podemos continuar este proceso y tendremos

$$a = p_1p_2 \ldots p_na_n$$

siempre que $a_1, a_2, \ldots, a_n$ no sean primos o iguales a ± 1; como

$$|a| > |a_1| \ldots > |a_{n-1}| > |a_n|$$

(puesto que un número primo p positivo es tal que $p \geq 2$); si no llegamos a a_i primo, llegaremos a $a_j = \pm 1$; luego

$$a = up_1p_2 \ldots p_n$$

donde $u = \pm 1$ y $p_1, p_2, \ldots, p_n$ son *enteros primos estrictamente positivos.* Demostremos que esta descomposición es *única,* sea

$$up_1p_2 \ldots p_n = vq_1q_2 \ldots q_m$$

desde luego $u = v$; ahora bien, p_1 primo divide uno de los factores del segundo miembro (corolario 1 del teorema 8) sea q_1, pero q_1 es igualmente primo positivo $p_1 = q_1$, se simplifica por p_1; haciendo este razonamiento un número finito de veces se llegará a agotar todos los factores de uno de los miembros, de donde si $r_1, r_2, \ldots, r_l$ son los factores primos restantes

$$r_1, r_2, \ldots, r_l = 1$$

igualdad imposible, pues $r_1, \ldots, r_l$ son superiores o iguales a 2; luego se agotan al mismo tiempo los factores de los dos miembros.

Finalmente, el razonamiento precedente no supone $p_1, \ldots, p_n$ distintos; reagrupando los factores iguales se obtiene:

TEOREMA 11. — *Todo número entero racional no nulo puede escribirse de una manera única bajo la forma*

$$a = u(p_1)^{k_1} \ldots (p_m)^{k_m}$$

donde $u = \pm 1$ *y* $p_1 \ldots p_m$ *son números enteros positivos primos todos distintos, y* $k_1 \ldots k_m$ *números enteros estrictamente positivos.*

APLICACIÓN 1: *Divisores de un entero.* — Los divisores de un entero a que tiene la descomposición anterior son todos de la forma

$$d = u(p_1)^{h_1} \ldots (p_m)^{h_m} \qquad u = \pm 1 \qquad 0 \leq h_i \leq k_i.$$

APLICACIÓN 2: m. c. d. *y* m. c. m. *de dos enteros descompuestos en factores primos.* — Designemos por $p_1, \ldots, p_n$ el *conjunto* de los factores primos supuestos positivos de a y de b, se puede escribir

$$a = u(p_1)^{k_1} \ldots (p_n)^{k_n} \qquad k_i \geq 0$$
$$b = v(p_1)^{l_1} \ldots (p_n)^{l_n} \qquad l_i \geq 0$$

naturalmente uno al menos de los enteros k_i y uno al menos de los enteros l_i no son nulos, se obtiene

$$D(a, b) = \prod_{i=1}^{i=n} (p_i)^{\inf(k_i, l_i)} \qquad M(a, b) = \prod_{i=1}^{i=n} (p_i)^{\sup(k_i, l_i)}.$$

EJERCICIOS

1. ¿Cuál es el número de divisores estrictamente positivos de un entero a estrictamente positivo? (Descomponer a en factores primos y utilizar la aplicación 1 anterior.)
2. Demostrar que la sucesión de enteros primos positivos es infinita (considerar el número $n! + 1$).

103. Observaciones sobre la divisibilidad en los anillos

a) La simplicidad de la teoría de la divisibilidad en **Z** proviene del hecho que **Z** es un anillo principal.

En todo anillo principal A (es decir, unitario, íntegro y en el que todo ideal es principal) todos los resultados de los §§ 98 al 102 que conciernen a **Z** permanecen válidos, excepto el ejercicio 2 del § 99 (algoritmo de Euclides), con aproximadamente el mismo vocabulario: es preciso reemplazar ± 1 por un elemento inversible cualquiera.

Además, a un ideal (a) no se le puede hacer corresponder, en general, un elemento único privilegiado (tal como $a > 0$ en **Z**), surge una gran complicación en los enunciados. Esta teoría es el objeto de los ejercicios 105 y 106.

b) Existen anillos principales particulares llamados *euclídeos,* provistos de una división euclídea (ver ej. 98) análoga a la de **Z**; veremos un importante ejemplo con el anillo de los polinomios en x de coeficientes en un cuerpo conmutativo (capítulo 11).

c) En el anillo **Z**, hemos dado *tres características equivalentes* de la noción de *entero primo*:

1. p es extremal (definición 2).
2. El ideal (p) es maximal (teorema 3).
3. $p \neq 0$, p es no inversible y el ideal (p) es primo (teorema 8).

Igualmente hemos dado en **Z** *tres características equivalentes* de la noción de enteros a y b *extraños*:

4. a y b sólo tienen como divisores comunes elementos inversibles (o su m. c. d. es 1) (definición 3 y teorema 5).
5. Igualdad de Bezout (teorema 5).
6. Para todo x, a divide bx implica que a divide x (o b divide ax implica que b divide x) (teorema 7).

Estas equivalencias son válidas en un anillo principal (ver ej. 105). Pero en un anillo no principal estas caracterizaciones pueden llevarnos a nociones distintas (ver ej. 109, 110 y 111 al final del capítulo), en particular en los anillos de polinomios en x e y volveremos a encontrar otros ejemplos (ver capítulo 11, § 195, y ej. 365).

IV. Cuerpos. Cuerpo Q de los racionales

104. Definiciones y propiedades generales

a) Definición. — *Un conjunto* K *provisto de una adición y de una multiplicación posee una* estructura de cuerpo *para esas dos operaciones si:*

1. K *posee una estructura de anillo para esas dos operaciones.*
2. $K^* = K - \{0\}$ (0 *elemento neutro de la adición*) *posee una estructura de grupo para la multiplicación.*

Se dirá que K es un *cuerpo para la adición y la multiplicación consideradas,* o simplemente un *cuerpo* si no es posible que surja confusión.

K* se llama el *grupo multiplicativo del cuerpo,* admite un elemento neutro $e \neq 0$, llamado *elemento unidad* del cuerpo: un cuerpo contiene al menos, pues, *dos elementos*: 0 y e.

Los axiomas de la estructura de cuerpo son, por lo tanto, distinguiendo los axiomas relativos a cada operación y los axiomas de "compatibilidad" entre esas dos operaciones (ver § 68),

$$\begin{cases} K_1 & (\forall a, b, c \in K) & (a+b)+c = a+(b+c) & \\ K_2 & (\exists 0 \in K)(\forall a \in K) & a+0 = 0+a = a & \\ K_3 & (\forall a \in K)(\exists a' \in K) & a+(a') = (a')+a = 0 & a' = -a \\ K_4 & (\forall a, b \in K) & a+b = b+a & \end{cases}$$

$$\begin{cases} K_5 & (\forall a, b, c \in K) & (ab)c = a(bc) & \\ K_6^{(17)} & (\exists e \in K)(\forall a \in K) & ae = ea = a & \\ K_7 & (\forall a \in K^*)(\exists a'' \in K^*) & aa'' = a''a = e & a'' = a^{-1} \end{cases}$$

$$\begin{cases} K_8 & (\forall a, b, c \in K) & a(b+c) = ab+ac \\ K_9 & (\forall a, b, c \in K) & (b+c)a = ba+ca. \end{cases}$$

Si la multiplicación es conmutativa, se dice que el cuerpo es *conmutativo.* Dos elementos particulares tales que $ab = ba$ se llaman permutables, el conjunto de los elementos permutables con todos los elementos del cuerpo es el *centro* del cuerpo.

Se llama característica de un cuerpo K, la *característica* de K considerada como anillo (ver § 97).

Todas las reglas de cálculo válidas en un anillo son válidas para un cuerpo. Además, todo elemento no nulo, siendo inversible, es regular para la multiplicación, luego: *un cuerpo no posee divisores de cero.* Además, para todo a de K y todo a' de K*

$$xa' = a \Leftrightarrow x = aa'^{-1}, \qquad a'y = a \Leftrightarrow y = a'^{-1}a$$

estos cocientes, por la derecha y por la izquierda (ver § 49), son en general distintos.

Si K es *conmutativo* $x = y = aa'^{-1}$, que se representa a menudo a/a'. Para todo $z \neq 0$ se tiene $a/a' = az/a'z$. Se verificará fácilmente las fórmulas siguientes $(a'b' \neq 0)$

$$\frac{a}{a'} + \frac{b}{b'} = \frac{ab' + ba'}{a'b'}, \qquad \frac{a}{a'}\frac{b}{b'} = \frac{ab}{a'b'}.$$

Finalmente en un cuerpo *conmutativo* la fórmula del binomio es siempre válida.

---

(17) La definición supone $ae = ea = a$ en K*; pero como en todo anillo unitario se tiene de $e0 = 0$, es verdadera en K.

b) EJEMPLOS Y EJERCICIOS

1. **Q**, **R** y **C** son los cuerpos conmutativos de característica nula.
2. Demostrar que existe un cuerpo de dos elementos (que son forzosamente el elemento cero y el elemento unidad, 0 y e), este cuerpo está definido por $e + e = 0$; es conmutativo y de característica 2.
3. $\mathbf{Z}/p\mathbf{Z}$ es un cuerpo conmutativo si y sólo si p es primo ($\dot{x} \neq \dot{0}$) demuestra que x no es un múltiplo de p, luego (§ 99) siendo p primo, x y p son primos entre sí; no existen, pues, los enteros x' y p' tales que

$$xx' + pp' = 1 \Rightarrow \dot{x}\dot{x}' = \dot{1}.$$

Por otro lado, si p no es primo, $\mathbf{Z}/p\mathbf{Z}$ (que contiene divisores de cero) no puede ser un cuerpo. Es de característica p. Deducir de lo anterior una nueva demostración del *teorema de* FERMAT (§ 97, ej. 2) *considerando el grupo* $(\mathbf{Z}/p\mathbf{Z})^*$.

4. De una manera más general todo anillo de integridad A unitario finito es un cuerpo (dado $a \neq 0$ la aplicación de A* en A* definida por $x \to ax$ es inyectiva, luego suprayectiva (§ 31); existe, pues, x' de A* tal que $aa' = e$). Se puede suponer simplemente A finito, unitario y sin divisores de cero (ver ej. 66, fin del capítulo 4).
5. El conjunto descrito por $a + a'\sqrt{2}$, a y a' describiendo Q es un cuerpo conmutativo.

105. Subcuerpo

DEFINICIÓN. — *Se llama* subcuerpo *de un cuerpo* K *toda parte no vacía* L *de* K, *estable respecto a las leyes de* K *y tal que la estructura inducida sobre* L *por estas leyes sea una estructura de cuerpo. Se dice que* K *es un* supercuerpo *o una* extensión *del cuerpo* L.

Se llama *subanillo* A *de un cuerpo* K todo subanillo de K considerado como anillo (por ejemplo, **Z** es un subanillo de **Q**). Se dice también que K es un supercuerpo del anillo A.

K es el mayor subcuerpo de K; todo subcuerpo de K distinto de K se llama *subcuerpo propio* de K. Se dice que un cuerpo es *primo* si no contiene otro subcuerpo que el mismo.

TEOREMA. — *Para que una parte no vacía* L *de un cuerpo* K *sea un subcuerpo de* K *es necesario y suficiente que:*

(1) $$(a \in \mathrm{L} \ \text{ y } \ b \in \mathrm{L}) \Rightarrow (a - b \in \mathrm{L} \ \text{ y } \ ab \in \mathrm{L})$$
(2) $$a \in \mathrm{L}^* \Rightarrow a^{-1} \in \mathrm{L}^*.$$

En efecto, según (1) (§ 93), L es un subanillo de K.

Por otra parte, L* es una parte estable de K para la multiplicación en K, luego la multiplicación inducida sobre L* es asociativa. (2) y la segunda parte de (1) muestran que para todo a de L*, $a^{-1}a = e$ pertenece a L*; luego con (2) vemos que L* posee las tres propiedades características de una estructura de grupo.

Se ve fácilmente que la *intersección* de una familia cualquiera de subcuerpos es también un subcuerpo de K. En particular, la intersección de todos los subcuerpos de K que contienen una parte X no vacía de K es un subcuerpo de K; es el menor subcuerpo que contiene X, se dice que está *engendrado* por X.

OBSERVACION

Puede suceder que un subanillo B de un anillo A tenga una estructura de cuerpo (ver § 146, ej. 5; § 158, ej. 4; cap. 8, ej. 200).

EJEMPLOS Y EJERCICIOS

1. **Q** es un subcuerpo de **R** y de **C**, **R** un subcuerpo de **C**.
2. Siendo p primo $\mathbf{Z}/p\mathbf{Z}$ es un cuerpo primo.
3. El centro de un cuerpo K es un subcuerpo de K. Generalmente el conjunto descrito por los elementos de K permutables con cada uno de los elementos de una parte X de un cuerpo K es un subcuerpo de K.
4. El cuerpo estudiado en el ejercicio 5, § 104, es un subcuerpo de **R**.
5. Demostrar que P intersección de todos los subcuerpos de K es primo y que es el único; demostrar que P está engendrado por e: se le llama el *subcuerpo primo* de K (ver § 107, d, ejercicio).

106. Ideales de un cuerpo. Homomorfismos. Isomorfismos de los cuerpos

a) Un cuerpo K teniendo una estructura de anillo para la suma y la multiplicación definidas sobre K, todas las definiciones y teoremas enunciados en los §§ 95 y 96 son válidas para los cuerpos, pero la teoría se simplifica bastante por el teorema siguiente:

TEOREMA. — *Los únicos ideales por la izquierda (resp. por la derecha) de un cuerpo* K, *considerado como anillo, son* $\{0\}$ *y* K.

En efecto, sea I un ideal, por la izquierda, por ejemplo, si $\mathrm{I} \neq \{0\}$ tenemos $a \neq 0$ perteneciente a I, pero existe entonces a^{-1} en K y $a^{-1}a = e$ pertenece a I y cualquiera que sea x de K, $xe = x$ pertenece a I, luego $\mathrm{I} = \mathrm{K}$; se demostraría igualmente que todo ideal por la derecha $\mathrm{I} \neq 0$ es idéntico a K.

EJERCICIO

Demostrar recíprocamente que todo anillo unitario A que sólo tiene como ideales los $\{0\}$ y A es un cuerpo ($a \neq 0$, se demostrará que $x \to \gamma_a(x) = ax$ y $x \to \delta_a(x)$ son suprayectivas y si A no es conmutativo se utilizará el ej. 66 del cap. 4).

¿Qué relación tiene este ejercicio con el ejercicio 4 del § 104?

b) Consideremos un homomorfismo f de un cuerpo K en un cuerpo K′, tenemos

$$(\forall\, x,\ y \in \mathrm{K}) \qquad f(x+y) = f(x) + f(y), \qquad f(xy) = f(x)f(y).$$

La propiedad de los homomorfismos de anillos enunciados en el § 96 (t. 2) y el teorema precedente muestran que $\mathrm{N} = f^{-1}(0)$ no puede ser otro que $\{0\}$ o bien K.

1. Si $\mathrm{N} = f^{-1}(0) = \{0\}$, la relación de equivalencia $x - y \in \mathrm{N}$ es la igualdad, luego el homomorfismo canónico de K sobre K/N definido por $x \to \dot{x}$ es

una biyección, y es también un *isomorfismo*; descomponiendo canónicamente f (ver § 96)

$$K \xrightarrow{s} K/N \xrightarrow{b} f(K) \xrightarrow{i} K'$$

b es un isomorfismo, y análogamente para $b \circ s$, luego $f(K)$ parte estable de K′ es isomorfo a K y tiene entonces una estructura de cuerpo: es un subcuerpo de K′.

2. Si $N = K$, la relación de equivalencia $x - y \in N$ es la equivalencia absoluta K/N tiene un solo elemento, así como $f(K)$; este único elemento de $f(K)$ es $f(0) = 0$, luego:

TEOREMA. — *Dado un homomorfismo f de un cuerpo K en un cuerpo K′, o bien $f(K)$ es un anillo reducido a cero, o bien $f(K)$ es un cuerpo y f es un isomorfismo de K sobre $f(K)$.*

Si f es suprayectivo, $K' = f(K)$ comprende al menos dos elementos 0 y e', y sólo el primer caso es posible.

COROLARIO. — *Todo homomorfismo suprayectivo de un cuerpo K sobre un cuerpo K′ es un isomorfismo de cuerpo.*

OBSERVACION

Se puede suponer que K′ es un conjunto provisto de una adición y de una multiplicación $f(K)$ es o bien el anillo cero o un cuerpo isomorfo a K.

107. Cuerpo de las fracciones de un anillo de integridad. Cuerpo Q de los racionales

a) Enunciado del problema de la inmersión de un anillo de integridad en un cuerpo conmutativo

En el § 59 hemos intentado sumergir un conjunto E, provisto de una ley interna satisfaciendo ciertas condiciones, en un grupo G de manera que E sea una parte estable de G; dado un anillo A, buscamos un cuerpo K tal que A sea un subanillo de K; como todo supercuerpo de K responde a la pregunta buscaremos K *minimal* (para la inclusión de los conjuntos).

Por otro lado, A, subanillo del cuerpo K, no deberá contener divisores de cero; en fin, por no ser el cero de A regular para la multiplicación, el conjunto A^* deberá estar sumergido en el grupo multiplicativo K^*. Hemos visto en el § 59 que este problema era posible si A^* (en consecuencia, A) y K^* eran conmutativos y si se modificaba ligeramente el enunciado primitivo: A^* debe de ser un grupo isomorfo a una parte estable de K^* para la multiplicación. Nos vemos así conducidos a formular el *problema de la inmersión* de un anillo de integridad A en un cuerpo conmutativo minimal K:

Dado un anillo de integridad A, *encontrar un cuerpo conmutativo minimal* K *tal que un subanillo* A′ *de* K *sea isomorfo a* A.

Vamos a proceder de la siguiente manera:

— K debe contener al menos el simetrizado (ver § 59) de A^* para la multiplicación, sea $\overline{A^*}$ este simetrizado.

— K debe contener también un elemento cero, sea ε, luego contener $A^* \cup \{\varepsilon\} = K'$.

— Si podemos extender la multiplicación de $\overline{A^*}$ a K′ y definir en K′ una adición tal que K′ sea un cuerpo, A al ser isomorfo a una parte A′ de K′, este conjunto K′ será una de las soluciones minimales buscadas.

— De hecho, observando que toda solución isomorfa a una solución es también una solución de nuestro problema, vamos a construir un cuerpo isomorfo a K′.

Demostraremos, en fin, que, salvo un isomorfismo, la solución del problema de la inmersión de un anillo de integridad en un cuerpo es único.

b) **Determinación de una solución**

Consideremos el simetrizado de A^* para la multiplicación, sea $\overline{A^*} = (A^* \times A^*)/R$, siendo R la relación de equivalencia $ab' = ba'$ entre elementos (a, a') y (b, b') de $A^* \times A^*$. En $\overline{A^*}$ la multiplicación está definida por

$$\left(\dot{\overline{a, a'}}\right)\left(\dot{\overline{b, b'}}\right) = \left(\dot{\overline{ab, a'b'}}\right)$$

existe un elemento unidad $\left(\dot{\overline{x, x}}\right)$, o $\left(\dot{\overline{e, e}}\right)$ si A es unitario, la inversa de $\left(\dot{\overline{a, a'}}\right)$ es $\left(\dot{\overline{a', a}}\right)$ y la aplicación f definida por

$$a \rightarrow f(a) = \left(\dot{\overline{ax, x}}\right)$$

(x elemento cualquiera de A^*) es un *isomorfismo* de A^* *sobre* $f(A^*)$ parte estable de $\overline{A^*}$ para la multiplicación.

En lo que sigue supondremos que A es unitario, lo que no supone restricción alguna (ver observación 2 más abajo); la aplicación f es tal que $f(a) = \left(\dot{\overline{a, e}}\right)$. Si ε es el elemento cero en $K' = \overline{A^*} \cup \{\varepsilon\}$ se deberá tener $\varepsilon^2 = \varepsilon$ y para todo $\left(\dot{\overline{x, x'}}\right)$ de $\overline{A^*}$: $\left(\dot{\overline{x, x'}}\right)\varepsilon = \varepsilon$. Por otra parte, para todo isomorfismo g de A sobre una parte de K′ se deberá tener: $g(0) = \varepsilon$. En lugar de considerar $\overline{A^*} = (A^* \times A^*)/R$ consideraremos

$$\overline{\overline{A}} = \frac{A \times A^*}{R_1},$$

siendo R_1 la relación de equivalencia $ab' = ba'$ entre los elementos (a, a') y (b, b') de $A \times A^*$, que induce R sobre A^*. Este nuevo conjunto cociente contiene $\overline{A}^*$ y el elemento $\left(\overset{\cdot}{\widehat{0, e}}\right)$.

Pero para todo $\left(\overset{\cdot}{\widehat{x, x'}}\right)$ de A

$$\left(\overset{\cdot}{\widehat{0, e}}\right)\left(\overset{\cdot}{\widehat{x, x'}}\right) = \left(\overset{\cdot}{\widehat{0x, ex'}}\right) = \left(\overset{\cdot}{\widehat{0, e}}\right)$$

y se ve fácilmente que la aplicación suprayectiva de A sobre $g(A) \subset \overline{A}$ definida por

$$a \to g(a) = \left(\overset{\cdot}{\widehat{a, e}}\right)$$

admite f por restricción a A^*, y es tal que

$$g(0) = \left(\overset{\cdot}{\widehat{0, e}}\right).$$

Se ve igualmente que $g(A)$ es una parte estable de $\overline{A}$ y que g es un isomorfismo de A y de $g(A)$ provistos de sus multiplicaciones respectivas.

Vamos a mostrar ahora que se puede dotar $\overline{A}$ de una *adición* tal que $\overline{A}$ sea un cuerpo conmutativo y que $g(A)$ sea un subanillo de $\overline{A}$, isomorfo a A. Esta última condición implicará que

$$\text{(1)} \qquad \left(\overset{\cdot}{\widehat{a, e}}\right) + \left(\overset{\cdot}{\widehat{b, e}}\right) = \left(\overset{\cdot}{\widehat{a + b, e}}\right).$$

Por otro lado, la multiplicación deberá ser distributiva respecto a la adición; luego, en particular (teniendo en cuenta (1)), deberemos tener

$$\left(\overset{\cdot}{\widehat{a'b', e}}\right)\left[\left(\overset{\cdot}{\widehat{a, a'}}\right) + \left(\overset{\cdot}{\widehat{b, b'}}\right)\right] = \left(\overset{\cdot}{\widehat{a'b', e}}\right)\left(\overset{\cdot}{\widehat{a, a'}}\right) + \left(\overset{\cdot}{\widehat{a'b', e}}\right)\left(\overset{\cdot}{\widehat{b, b'}}\right)$$

$$= \left(\overset{\cdot}{\widehat{ab', e}}\right) + \left(\overset{\cdot}{\widehat{ba', e}}\right) = \left(\overset{\cdot}{\widehat{ab' + ba', e}}\right)$$

en consecuencia ($a'b' \neq 0$, perteneciendo a' y b' a A^*),

$$\left(\overset{\cdot}{\widehat{a, a'}}\right) + \left(\overset{\cdot}{\widehat{b, b'}}\right) = \left(\overset{\cdot}{\widehat{ab' + ba', e}}\right)\left[\left(\overset{\cdot}{\widehat{a'b', e}}\right)\right]^{-1} = \left(\overset{\cdot}{\widehat{ab' + ba', e}}\right)\left(\overset{\cdot}{\widehat{e, a'b'}}\right)$$

$$\text{(2)} \qquad \left(\overset{\cdot}{\widehat{a, a'}}\right) + \left(\overset{\cdot}{\widehat{b, b'}}\right) = \left(\overset{\cdot}{\widehat{ab' + ba', a'b'}}\right)$$

La fórmula (2) sólo será válida si se demuestra que:

1. El resultado de la operación definida por (2) es independiente de los representantes escogidos.

2. La suma definida en (2) es distributiva en relación con la multiplicación.

3. $\overline{\mathrm{A}}$ es un cuerpo en el que la parte descrita por $\left(\overset{\cdot}{\widehat{a, e}}\right)$ es un subanillo isomorfo de A.

La primera parte se verifica observando que

$$\left(\overset{\cdot}{\widehat{a, a'}}\right) = \left(\overset{\cdot}{\widehat{x, x'}}\right) \Leftrightarrow ax' = a'x$$
$$\left(\overset{\cdot}{\widehat{b, b'}}\right) = \left(\overset{\cdot}{\widehat{y, y'}}\right) \Leftrightarrow by' = b'y$$

un cálculo fácil si se utiliza el hecho de que $a'b'x'y' \neq 0$, demuestra que

$$\left(\overset{\cdot}{\widehat{ab' + ba',\ a'b'}}\right) = \left(\overset{\cdot}{\widehat{xy' + yx',\ x'y'}}\right).$$

La distributividad se verifica igualmente

$$\left(\overset{\cdot}{\widehat{c, c'}}\right)\left[\left(\overset{\cdot}{\widehat{a, a'}}\right) + \left(\overset{\cdot}{\widehat{b, b'}}\right)\right] = \left(\overset{\cdot}{\widehat{c, c'}}\right)\left(\overset{\cdot}{\widehat{ab' + ba',\ a'b'}}\right) =$$
$$\left[\overset{\cdot}{\widehat{c(ab' + ba'),\ c'a'b'}}\right] = \left(\overset{\cdot}{\widehat{cab',\ c'a'b'}}\right) + \left(\overset{\cdot}{\widehat{cba',\ c'a'b'}}\right) =$$
$$\left(\overset{\cdot}{\widehat{c, c'}}\right)\left(\overset{\cdot}{\widehat{a, a'}}\right) + \left(\overset{\cdot}{\widehat{c, c'}}\right)\left(\overset{\cdot}{\widehat{b, b'}}\right).$$

En fin, $\overline{\mathrm{A}}$ es un grupo abeliano para la adición: se verificará mediante un cálculo pesado, pero fácil, que la adición (2) es *asociativa,* que $\left(\overset{\cdot}{\widehat{0, e}}\right)$ es el *elemento cero* y que todo elemento $\left(\overset{\cdot}{\widehat{a, a'}}\right)$ tiene un *opuesto* $\left(\overset{\cdot}{\widehat{-a, a'}}\right)$.

$(\overline{\mathrm{A}})^*$, que no es otro si no el simétrico $\overline{\mathrm{A}^*}$ de A^*, es un grupo multiplicativo abeliano y $\left(\overset{\cdot}{\widehat{a, a'}}\right)\left(\overset{\cdot}{\widehat{0, e}}\right) = \left(\overset{\cdot}{\widehat{0, e}}\right)\left(\overset{\cdot}{\widehat{a, a'}}\right) = \left(\overset{\cdot}{\widehat{0, e}}\right)$; por otra parte, la multiplicación es distributiva respecto a la suma; en consecuencia, $\overline{\mathrm{A}}$ es un *cuerpo*; siendo igual a $\overline{\mathrm{A}^*} \cup \left\{\left(\overset{\cdot}{\widehat{0, e}}\right)\right\}$ es *minimal*. En fin, la aplicación g definida por

$$a \to g(a) = \left(\overset{\cdot}{\widehat{a, e}}\right)$$

es un isomorfismo de A sobre g(A): es, en efecto, una biyección y para todo par $\left(\overset{\cdot}{\widehat{a, e}}\right)$, $\left(\overset{\cdot}{\widehat{b, e}}\right)$, de elementos de g(A) se tiene fácilmente (según las propiedades de las operaciones en $\overline{\mathrm{A}}$)

$$g(a) + g(b) = \left(\overset{\cdot}{\widehat{a,\, e}}\right) + \left(\overset{\cdot}{\widehat{b,\, e}}\right) = \left(\overset{\cdot}{\widehat{a + b,\, e}}\right) = g(a + b)$$

$$g(a)g(b) = \left(\overset{\cdot}{\widehat{a,\, e}}\right)\left(\overset{\cdot}{\widehat{b,\, e}}\right) = \left(\overset{\cdot}{\widehat{ab,\, e}}\right) = g(ab).$$

El cuerpo $\overline{A} = (A \times A^*)/R_1$ *así construido responde completamente a la pregunta.*

c) **Unicidad de la solución** (salvo un isomorfismo)

Sea K un cuerpo que responde a la pregunta, designemos por A′ la parte de K isomorfa al anillo de integridad dado A; A′ es, en consecuencia, también un anillo de integridad, ya que hemos supuesto que A tenía un elemento unidad *e*, asimismo lo tiene A′: su elemento unidad e' es el de K.

Designando también por R_1 la relación $ab' = ba'$, definida sobre $A' \times A'^*$, pongamos $\overline{A'} = (A' \times A'^*)/R_1$; es claro que $\overline{A}$ y $\overline{A'}$ son cuerpos conmutativos isomorfos.

Ahora bien, la definición de un elemento $\left(\overset{\cdot}{\widehat{a,\, a'}}\right)$ de $\overline{A'}$ demuestra que a este elemento se le puede asociar un *único* elemento de K poniendo

$$\varphi\left(\left(\overset{\cdot}{\widehat{a,\, a'}}\right)\right) = aa'^{-1}.$$

Las propiedades de las operaciones en $\overline{A'}$ (que son las mismas que las de $\overline{A}$) y las de las operaciones en el cuerpo conmutativo K (ver § 104) muestran que la aplicación así definida φ de $\overline{A'}$ en K es un homomorfismo de cuerpos; según esto $\varphi\left(\left(\overset{\cdot}{\widehat{e',\, e'}}\right)\right) = e' \neq 0$, luego (§ 106, *b*) $\varphi(\overline{A'})$ es un cuerpo isomorfo de $\overline{A'}$; en consecuencia, a $\overline{A}$ y responde a la pregunta. De ello resulta que $\varphi(\overline{A'}) = K$, si no K no sería un cuerpo minimal respondiendo a la pregunta, de donde:

Teorema. — *Dado un anillo de integridad* A, *hay un cuerpo conmutativo minimal* K, *único a un isomorfismo, tal que un subanillo de* K *sea isomorfo a* A.

Se identifica todos estos cuerpos a $\overline{A}$. Identificando, para todo x de A, $\left(\overset{\cdot}{\widehat{x,\, e}}\right)$ y x tendremos ($a' \neq 0$)

$$a'\left(\overset{\cdot}{\widehat{a,\, a'}}\right) = \left(\overset{\cdot}{\widehat{a',\, e}}\right)\left(\overset{\cdot}{\widehat{a,\, a'}}\right) = \left(\overset{\cdot}{\widehat{a'a,\, a'e}}\right) = \left(\overset{\cdot}{\widehat{a,\, e}}\right) = a$$

en consecuencia, $\left(\overset{\cdot}{\widehat{a,\, a'}}\right)$ es el *cociente* aa'^{-1}. Se le *representa* por la fracción

a/a' o por *cualquier* otra fracción b/b' tal que $ab' = ba'$; luego la *fracción* a/a' no es más que un *representante* del elemento $\left(\overset{\cdot}{\widehat{a,\ a'}}\right)$ de $\overline{\mathrm{A}}$; es, pues, a causa de un abuso de lenguaje que se dice que $\overline{\mathrm{A}}$ *es el cuerpo de fracciones del anillo de integridad* A.

Si a y a' son extraños en A se dice que la fracción a/a' es *irreducible.*

En ciertos casos se podrá definir un representante privilegiado de $\left(\overset{\cdot}{\widehat{a,\ a'}}\right)$ (por ejemplo, en **Q** (ver más abajo) y en el cuerpo de las fracciones racionales en x, ver § 200, *a*).

OBSERVACIONES

1. Si se hubiera tenido sólo la intención de establecer la existencia de K, sin demostrar la unicidad, hubiera sido suficiente considerar $\overline{\mathrm{A}}$ y proporcionarle *a priori* la multiplicación y la adición *que hemos encontrado*, la exposición hubiera sido más simple, pero el resultado final menos interesante.

2. Se podrá demostrar a título de ejercicio que el resultado final subsiste si se supone A no unitario (para todo x no nulo de A, $\left(\overset{\cdot}{\widehat{x,\ x}}\right)$ es elemento unidad de $\overline{\mathrm{A}}$).

En fin, si $\mathrm{A} = \{0\}$, se puede siempre sumergirlo en el cuerpo de dos elementos (ver § 104, ej. 2).

d) Cuerpo Q de los números racionales

Aplicando la teoría precedente al anillo de integridad **Z**, obtenemos el cuerpo **Q** de los números racionales; cada uno de los elementos de **Q**, o sea, $\left(\overset{\cdot}{\widehat{a,\ a'}}\right)$ es una clase de equivalencia que admite como representante toda fracción de numerador x y de denominador $x \neq 0$ tal que $ax' = a'x$. Identificando $\left(\overset{\cdot}{\widehat{a,\ 1}}\right)$ de **Q** y a de **Z**, obtenemos

$$\mathbf{Z} \subset \mathbf{Q}.$$

Se escribe como siempre $\mathbf{Q}^* = \mathbf{Q} - \{0\}$.

En el caso de **Q** se puede definir un representante privilegiado de $\left(\overset{\cdot}{\widehat{a,\ a'}}\right)$ (ver § 18, *c*), es la fracción a/a' *irreducible*, es decir, tal que a y a' sean primos entre sí, de *denominador positivo*; en efecto, sea a/a' y b/b' dos fracciones que representan el mismo número racional

$$a/a' = b/b', \qquad \mathrm{D}(a,\ a') = \mathrm{D}(b,\ b') = 1, \qquad a' > 0, \qquad b' > 0$$

$ab' = ba'$ implica que a' y b' se dividen mutuamente, luego $b' = \pm a'$, como a' y b' son positivos $a' = b'$ y $a = b$.

EJERCICIO

Sea P el subcuerpo primo de un cuerpo K (§ 105, ej. 5) y sea p la característica de K. Demostrar que:

— o bien $p = 0$ y P es isomorfo a **Q**,
— o bien p es primo y P es isomorfo a $\mathbf{Z}/p\mathbf{Z}$.

(Observar que P contiene el *subanillo* E, de K, engendrado por e; considerar el homomorfismo f de **Z** sobre E definido por $f(n) = ne$; observar seguidamente que E es isomorfo a $\mathbf{Z}/p\mathbf{Z}$ y que un cuerpo no puede contener divisores de cero.)

V. Anillos y cuerpos ordenados. Nociones sobre el cuerpo R

108. Grupos, anillos, cuerpos ordenados. Orden en Q

a) Hemos indicado en los §§ 60 y 62 que la relación $a \leq b$ daba al grupo **Z** y al anillo **Z** una estructura de orden total, compatible, en un sentido que hemos indicado, con la estructura de grupo o de anillo de **Z**. Damos aquí las definiciones generales:

Definición 1. — *Un grupo abeliano* G, *con notación aditiva, provisto de una relación de orden* $a \leq b$ *es un* grupo ordenado *si*

$$(\forall x \in G) \qquad a \leq b \Rightarrow a + x \leq b + x.$$

Se dice entonces que las estructuras de grupo y de orden son compatibles. Si el orden es total se dice que G *es un* grupo totalmente ordenado.

Resulta inmediatamente de esta definición que en un grupo ordenado $a \geq 0$ implica $a - a = 0 \geq -a$; por consiguiente,

$$a \geq 0 \Leftrightarrow (-a) \leq 0.$$

EJERCICIOS

1. En un grupo ordenado las relaciones $(1 \leq i \leq n)$ $a_i \leq b_i$ implican

$$a_1 + a_2 + \ldots a_n \leq b_1 + b_2 + \ldots + b_n.$$

2. En un grupo ordenado $a \leq b$ es equivalente a $a + c \leq b + c$.

3. Siendo P una parte de un grupo abeliano G aditivo tal que $P + P \subset P$, demostrar que la relación $b - a \in P$ da a G una estructura de grupo ordenado si y solamente si $P \cap (-P) = \{0\}$. El orden es total si y sólo si $G = P \cup (-P)$ (ver ej. 88 y 89 final del capítulo 4).

4. Si G es un grupo ordenado, $G \times G$ provisto de una de las dos relaciones de orden:

a) $(a_1, a_2) \leq (b_1, b_2) \Leftrightarrow (a_1 \leq b_1, a_2 \leq b_2)$,
b) $(a_1, a_2) \leq (b_1, b_2) \Leftrightarrow (a_1 < b_1 \quad \text{o} \quad a_1 = b_1, a_2 \leq b_2)$

¿es un grupo ordenado?

5. Todo grupo monógeno totalmente ordenado no reducido a $\{0\}$ es infinito.

DEFINICIÓN 2. — *Un anillo (o un cuerpo) conmutativo* E *provisto de una relación de orden* $a \leq b$ *es un* anillo (o un cuerpo) ordenado *si*

$$(\forall x \in E) \qquad a \leq b \Rightarrow a + x \leq b + x$$
$$(a \geq 0,\ b \geq 0) \Rightarrow ab \geq 0.$$

Se dice entonces que las estructuras de anillo (o de cuerpo) y de orden son compatibles. Si el orden es total se dice que E *és un* anillo (o un cuerpo) totalmente ordenado.

Los elementos tales que $x \geq 0$ (resp. $x \leq 0$) de un *anillo* totalmente ordenado se llaman elementos *positivos* (resp. *negativos*) del anillo. En tales anillos *todo cuadrado es positivo.*

EJERCICIOS

6. Siendo P una parte de un anillo A, la relación $b - a \in P$ proporciona a A una estructura de anillo ordenado si y sólo si

$$P + P \subset P \qquad PP \subset P \qquad P \cap (-P) = \{0\}.$$

El orden es total si y sólo si

$$A = P \cup (-P).$$

7. Todo anillo totalmente ordenado es de característica nula (ver ej. 5).

b) **Q es un cuerpo totalmente ordenado**

Vamos a mostrar que se puede dar al cuerpo **Q** una relación de orden que se deduce de la ya definida sobre **Z**, que hace de **Q** un cuerpo totalmente ordenado y ello de una manera única.

Designemos por $\alpha \leq \beta$ esta relación de orden sobre **Q**, será compatible con la adición en **Q**, luego

$$\alpha > 0 \Rightarrow \alpha - \alpha > -\alpha \Leftrightarrow -\alpha < 0$$

además, razonando por inducción sobre el entero $n > 0$: $\alpha > 0$ implica $n\alpha > 0$. Por otro lado,

(1) $$n > 0 \Rightarrow 1/n > 0.$$

Si no $1/n < 0$ (el orden es total y $1/n \neq 0$); en consecuencia, $-1/n > 0$ y siendo positivo el producto de dos elementos positivos se tendría $n(-1/n) = -1 > 0$, lo que es incompatible con el hecho que la relación $\alpha \leq \beta$ induzca sobre **Z** la relación de orden que allí está ya definida.

Esta propiedad (1) comprende el hecho de que

$$(a > 0 \quad \text{y} \quad a' > 0) \Rightarrow \frac{a}{a'} = a\left(\frac{1}{a'}\right) > 0.$$

Designemos por $\mathbf{Q}_+$ el conjunto de los racionales positivos ($\alpha \geq 0$), es idéntico al conjunto de los racionales con un representante de la forma a/a'

con $a \geq 0$ y $a' > 0$. Toda relación de orden que responda a las condiciones impuestas será, en consecuencia, tal que

$$\alpha \leq \beta \Leftrightarrow \beta - \alpha \geq 0 \Leftrightarrow \beta - \alpha \in \mathbf{Q}_+.$$

Recíprocamente sea a/a' y b/b' representantes privilegiados de α y β (*fracciones irreducibles de denominador positivo*)·

$$\beta - \alpha = \frac{ba' - ab'}{a'b'} \in \mathbf{Q}_+ \Leftrightarrow ba' - ab' \geq 0$$

es decir,

$$\alpha \leq \beta \Leftrightarrow ab' \leq ba'$$

se verifica inmediatamente, gracias a las propiedades del anillo ordenado **Z**, que se trata de un orden total y que

$$\frac{a}{a'} \leq \frac{b}{b'} \Rightarrow \frac{a}{a'} + \frac{c}{c'} \leq \frac{b}{b'} + \frac{c}{c'}$$

$$\left(\frac{a}{a'} \geq 0 \quad \text{y} \quad \frac{b}{b'} \geq 0 \right) \Rightarrow \frac{ab}{a'b'} \geq 0$$

y finalmente que esta relación induce efectivamente sobre **Z** la relación $a \leq b$, luego:

TEOREMA. — *Existe sobre el cuerpo* **Q** *una sola relación de orden que induce sobre* **Z** *la relación* $a \leq b$ *y proporcionando a* **Q** *una estructura de cuerpo totalmente ordenado.*

Se designará por $\mathbf{Q}_+^*$ el conjunto de los *racionales estrictamente positivos*, es un grupo para la multiplicación, se ha obtenido ya en el § 59, *e*, como simétrico de $\mathbf{N}^*$ para la multiplicación.

Se designará por $\mathbf{Q}_-$ el conjunto de los *racionales negativos* ($\alpha \leq 0$) y por $\mathbf{Q}_-^*$ el conjunto de los *racionales estrictamente negativos* ($\alpha < 0$).

c) Propiedades del orden definido sobre Q

Sea $\alpha = a/a' > 0$ y $\beta = b/b' > 0$ ($a' > 0$, $b' > 0$) dos racionales, existe un entero $n > 0$ tal que $nab' > a'b$, pues el orden definido sobre **Z** es *arquimediano* (ver § 62); resulta de ello que existe $n > 0$ tal que $n\alpha > \beta$, de donde:

TEOREMA. — *El cuerpo* **Q** *de los racionales ordenado por la relación* $\alpha \leq \beta$ *es un cuerpo totalmente ordenado arquimediano, es decir, para todo par de racionales* α, β *estrictamente positivos existe un entero natural* n *estrictamente positivo tal que* $n\alpha > \beta$.

Dados dos racionales distintos $\alpha < \beta$, existe siempre al menos un racional γ tal que $\alpha < \gamma < \beta$; por ejemplo, $(\alpha + \beta)/2$, de donde:

TEOREMA. — *Todos los intervalos abiertos* $]\alpha, \beta[$ ($\alpha < \beta$) *de* **Q** *no son vacíos. Se dice que, para el orden,* **Q** *es denso sobre sí mismo.*

109. Cuerpo de los reales

a) Introducción. Enunciado del teorema de existencia

Consideremos una parte X acotada superiormente de **Q**, no existe en general *en* **Q** una más pequeña cota superior de X, es decir, un límite superior de X.

Por ejemplo (ver § 23, ej. 4), el conjunto X de los racionales x tales que $x \geq 0$ y $x^2 < 2$ no tiene límite superior en **Q**.

Observemos que si en un grupo aditivo totalmente ordenado toda parte X acotada superiormente admite un límite superior, resulta que toda parte acotada inferiormente Y admite un límite inferior (basta aplicar a $-$Y el resultado sobre el límite superior).

Nos vemos así conducidos a preguntarnos si existen supercuerpos E de **Q** totalmente ordenados, arquimedianos y tales que toda parte acotada superiormente de E tenga un límite superior en E: la respuesta es positiva, y dada por el siguiente teorema, que admitiremos sin demostración:

Teorema. — *Sea* E *un conjunto provisto de una adición, de una multiplicación y de una relación de orden, existen conjuntos* E *tales que:*

E_1. E *es un cuerpo conmutativo.*
E_2. E *es un cuerpo totalmente ordenado.*
E_3. *El orden definido sobre* E *es arquimediano.*
E_4. *El orden definido sobre* E *es tal que toda parte acotada superiormente de* E *admite un límite superior.*

Todos estos conjuntos son isomorfos para la estructura de cuerpo y para la estructura de orden.

Designaremos por **R** *aquel cuerpo* E *que tiene por elemento unidad, el elemento unidad* 1 *del cuerpo de los racionales; los elementos de* **R** *se llaman números reales.*

En consecuencia, **R**, *como grupo aditivo contiene* **Z**, *y como cuerpo contiene el cuerpo de las fracciones de* **Z**, o sea, **Q**. Todo elemento del complementario de **Q** con relación a **R** se llama número *irracional.*

Se designa por $\mathbf{R}^*$ el conjunto de los números reales no nulos (este conjunto tiene una estructura de grupo conmutativo para la multiplicación), por $\mathbf{R}_+$ (resp. $\mathbf{R}_-$) el conjunto de los números *reales positivos* (resp. *negativos*), por $\mathbf{R}^*_+$ (resp. $\mathbf{R}^*_-$) el conjunto de los números *reales estrictamente positivos* (resp. estrictamente *negativos*).

Se puede construir un tal cuerpo **R** de varias maneras considerando ciertos conjuntos de partes de **Q** que se dota de una estructura de grupo aditivo totalmente ordenado, luego de una multiplicación. Los dos procesos más clásicos utilizan:

1. Las sucesiones de Cauchy de los números racionales.

2. Las cortaduras de Dedekind (o proceso análogo: las secciones inferiormente abiertas de números racionales).

La construcción de **R** con la ayuda de las sucesiones de CAUCHY se expone en el tomo II de este curso[18] (Análisis). Los ejercicios del número 114 al 117 dan la construcción de **R** mediante las secciones inferiormente abiertas de **Q**

b) Reglas de cálculo en R[19]

Se obtienen de los axiomas E_1, E_2, E_3, E_4 enunciados anteriormente.

1. Siendo **R** un cuerpo conmutativo, son válidas las reglas dadas en los §§ 91, 92 y 104, en particular la fórmula del binomio.

R* es un grupo conmutativo para la multiplicación, luego cualquiera que sean los reales x e y *no nulos* y los enteros racionales m y n (tabla del § 70)

$$x^m x^n = x^{m+n}, \qquad (xy)^m = x^m y^m, \qquad (x^m)^n = x^{mn}.$$

2. Siendo **R** un cuerpo totalmente ordenado:

— es de *característica nula* (§ 108, ej. 7);

— $x \leq y$ y $z \geq 0 \Rightarrow xz \leq yz$;

— *todo cuadrado es positivo* (§ 108 *b*), luego

$$(x_1)^2 + (x_2)^2 + \dots + (x_n)^2 = 0 \Leftrightarrow (x_1 = x_2 = \dots = x_n = 0).$$

— En fin, la fórmula (*n entero estrictamente positivo*)

$$x^n - y^n = (x - y)(x^{n-1} + x^{n-2}y + \dots + xy^{n-2} + y^{n-1})$$

muestra que la aplicación $x \to x^n$ de $\mathbf{R}_+$ en $\mathbf{R}_+$ es *estrictamente creciente* para n entero estrictamente positivo.

3. Se demuestra en Análisis[20] que para $n \in \mathbf{N}^*$ la aplicación $x \to x^n$ de $\mathbf{R}_+$ en el mismo es suprayectiva; como es estrictamente creciente es ***biyectiva***, luego todo real positivo x tiene una *raíz n-ésima positiva única* representada $\sqrt[n]{x}$ (véase ej. 118).

El símbolo $\sqrt[n]{\ }$ (representado $\sqrt{\ }$ si $n = 2$) es un *radical de índice n*, estudiaremos el cálculo de los radicales en el § 111.

EJERCICIO

Sea f un endomorfismo del cuerpo **R**.

a) Se supone $f(1) \neq 0$. ¿Cuál es entonces el valor de $f(1)$? Deducir el valor $f(x)$ para x racional.

Demostrar seguidamente que f es estrictamente creciente (utilizar el hecho que todo real estrictamente positivo es un cuadrado).

b) Demostrar que f es bien la aplicación nula, bien la aplicación idéntica de **R** en **R**

(18) *N. del T.* — Se trata del libro de R. COUTY y J. EZRA *Analyse*, de la Editorial Armand Colin, de París.

(19) El estudio de las nociones aquí indicadas, así como las de los párrafos 110 y 111, se desarrollan en el curso de Análisis. Sólo retenemos los resultados que nos serán útiles en el curso de Algebra.

110. Valor absoluto y distancia en R. Aplicaciones

a) DEFINICIÓN 1. — *Se llama* valor absoluto *del número real x, el número real expresado con la notación* $|x|$ *y definido por*

$$|x| = \sup(x, -x).$$

Se tiene entonces que $|-x| = |x|$ y

$$x \geq 0 \Rightarrow |x| = x, \qquad x \leq 0 \Rightarrow |x| = -x$$

se ve fácilmente que la aplicación $x \to |x|$ de $\mathbf{R}$ en $\mathbf{R}_+$ verifica las propiedades siguientes

$$x = 0 \Leftrightarrow |x| = 0$$
$$(\forall x, y \in \mathbf{R}) \qquad |xy| = |x||y|, \qquad |x + y| \leq |x| + |y|.$$

De lo que se deduce

$$\big||x| - |y|\big| \leq |x + y| \leq |x| + |y|$$
$$\big||x| - |y|\big| \leq |x - y| \leq |x| + |y|$$
$$|x_1 + x_2 + \ldots + x_n| \leq |x_1| + |x_2| + \ldots + |x_n|.$$

DEFINICIÓN 2. — *Dado un cuerpo* K, *si existe una aplicación* ν *de* K *en* $\mathbf{R}_+$ *que verifica*

V_1. $x = 0 \Leftrightarrow \nu(x) = 0$
V_2. $(\forall x, y \in K) \qquad \nu(xy) = \nu(x)\nu(y)$
V_3. $(\forall x, y \in K) \qquad \nu(x + y) \leq \nu(x) + \nu(y)$

se dice que K *es un* cuerpo valorado; ν *se llama un* valor absoluto *definido sobre* K; *por abuso de lenguaje se dice que* $\nu(x)$, *a menudo expresado* $|x|$, *es un valor absoluto de* x.

Se ve que $\mathbf{R}$ y $\mathbf{Q}$ son cuerpos valorados.

La noción de valor absoluto nos permite precisar las propiedades de la aplicación $x \to x^n$ de $\mathbf{R}$ en $\mathbf{R}$ (n entero > 0). Sabemos que su restricción a $\mathbf{R}_+$, toma sus valores en $\mathbf{R}_+$, es biyectiva y estrictamente creciente; por otra parte, esta aplicación es par para n par e impar para n impar, se deduce de todo ello que

$n = 2p + 1 > 0$	$n = 2p > 0$
$x^{2p+1} = y^{2p+1} \Leftrightarrow x = y$	$x^{2p} = y^{2p} \Leftrightarrow \|x\| = \|y\|$
$x^{2p+1} < y^{2p+1} \Leftrightarrow x < y$	$x^{2p} < y^{2p} \Leftrightarrow \|x\| < \|y\|$

b) Consideremos la aplicación d de $\mathbf{R} \times \mathbf{R}$ en $\mathbf{R}_+$ definida por

$$d(x, y) = |x - y|$$

se ve fácilmente que d verifica

$$x = y \Leftrightarrow d(x, y) = 0$$
$$(\forall x, y \in \mathbf{R}) \qquad d(y, x) = d(x, y)$$
$$(\forall x, y, z \in \mathbf{R}) \qquad d(x, y) \leq d(x, z) + d(z, y).$$

DEFINICIÓN 3. — *Dado un conjunto* E, *toda aplicación* d *de* $E \times E$ *en* $\mathbf{R}_+$ *que verifique*

D_1. $x = y \Leftrightarrow d(x, y) = 0$
D_2. $(\forall x, y \in E) \qquad d(y, x) = d(x, y)$
D_3. $(\forall x, y, z \in E) \qquad d(x, y) \leq d(x, z) + d(z, y)$

se llama una distancia. *Se llama* espacio métrico *todo conjunto* E *provisto de una distancia.*

La tercera propiedad se conoce con el nombre de "*desigualdad triangular*" (ver § 118). Por *abuso de lenguaje* se dice que x e y se hallan a la distancia $d(x, y)$. Se ve que, provistos de $d(x, y) = |x - y|$, **R** y **Q** son espacios métricos; se les llama, respectivamente, *recta numérica* y *recta racional.*

EJERCICIOS

1. Demostrar que en $\mathbf{R}^n$ descrito por $x = (x_1, x_2, \ldots, x_n)$ las aplicaciones siguientes de $\mathbf{R}^n \times \mathbf{R}^n$ en $\mathbf{R}_+$ son distancias:

a) $d_1(x, y) = |x_1 - y_1| + \ldots + |x_n - y_n|$,

b) $d_2(x, y) = \sup |x_i - y_i| \qquad i \in [1, n]$,

c) $d_3(x, y) = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2}$ (distancia *euclídea*)

(se podrá empezar con el estudio del caso $n = 2$).

2. Demostrar que si d es una distancia definida sobre E, análogamente lo es $\delta = d/(1 + d)$.

111. Cálculos con radicales. Exponentes fraccionarios

a) Raíces n-ésimas de x

Las propiedades de la aplicación $x \to x^n$ $(n > 0)$ de **R** en **R** muestran que las soluciones de $y^n = x$, es decir, las *raíces* n-ésimas de x son, en **R** (n entero estrictamente positivo),

$$n \text{ par} \begin{cases} x > 0 & y' = +\sqrt[n]{x}, \quad y'' = -\sqrt[n]{x} \\ x < 0 & \text{no hay solución} \end{cases}$$

$$n \text{ impar} \begin{cases} x > 0 & y = \sqrt[n]{x} \\ x < 0 & y = -\sqrt[n]{-x} \end{cases}$$

para todo $n > 0$: $x = 0, \quad y = 0.$

b) Cálculos con los radicales

Observemos seguidamente que $\sqrt[n]{x}$, así como x, son positivos por definición en particular

$$\sqrt{x^2} = |x|.$$

Utilizando la definición de $\sqrt[n]{x}$ y las fórmulas del § 109, *b*, se demostrará fácilmente que, cualesquiera que sean los números reales x e y positivos, y los enteros naturales m y n estrictamente positivos,

$$\sqrt[m]{x} = \sqrt[mn]{x^n}, \qquad \sqrt[m]{xy} = \sqrt[m]{x}\sqrt[m]{y}, \qquad (\sqrt[m]{x})^n = \sqrt[m]{x^n}.$$

OBSERVACION

Estas fórmulas pueden ser inexactas si x o y son negativos; por ejemplo, si se designa la raíz cúbica de -8 con la notación $\sqrt[3]{-8} = -2$ (*notación incorrecta* según la definición de un radical), se tiene $-2 = \sqrt[3]{-8} \neq \sqrt[6]{(-8)^2} = 2$.

c) Exponentes fraccionarios

Si x es un número real *positivo*, la primera fórmula anterior sobre los radicales muestra que (con p, q, p', q' enteros estrictamente positivos)

$$\frac{p}{q} = \frac{p'}{q'} \qquad \sqrt[q]{x^{p'}} = \sqrt[q']{x^{p'}}.$$

Este número positivo único es, en consecuencia, asociado al *número racional* pq^{-1} y no a la fracción p/q. Por otro lado, si pq^{-1} es igual al entero natural n: $\sqrt[q]{x^p} = x^n$; resulta que podemos poner, sin peligro de contradicción, si p y q son estrictamente positivos, así como x,

$$x^{(p/q)} = \sqrt[q]{x^p}, \qquad x^{-(p/q)} = [x^{(p/q)}]^{-1}.$$

Se verificará fácilmente, con la ayuda de las fórmulas precedentes, y teniendo en cuenta la igualdad $x^0 = 1$ $(x \neq 0)$, que cualquiera que sean los reales x e y *estrictamente positivos* y los racionales r y s, se tiene

$$x^r x^s = x^{r+s}, \qquad (xy)^r = x^r y^r, \qquad (x^r)^s = x^{rs}.$$

Ejercicios

90. Se considera un anillo A en el que todo elemento es independiente para la multiplicación, es decir, para todo x de A, $x^2 = x$ (cap. 3, ej. 51). Este anillo se llama anillo de *Boole*.

a) Demostrar que A es de característica 2 y que A es conmutativo (escribir que $x + x$ y $x + y$ son idempotentes).

b) Demostrar que para todo par de elementos x, y de A $xy(x + y) = 0$. Deducir que si A es íntegro se reduce a $\{0\}$ o es isomorfo a $\mathbf{Z}/2\mathbf{Z}$ y que, si A está provisto de divisores de cero, posee al menos tres elementos.

c) Demostrar de una manera más precisa que todo anillo de *Boole*, no reducido a cero, tiene dos elementos o al menos cuatro. Demostrar que sólo hay una estructura de anillo de *Boole* con cuatro elementos.

d) Demostrar que $\mathfrak{F}(E)$ provisto de la «adición» $(A, B) \to A \Delta B$ (diferencia simétrica; ver § 5, ej. 1) y de la «multiplicación» $(A, B) \to A \cap B$ es un anillo de *Boole*.

91. Sea G un grupo abeliano (designado aditivamente), se considera el conjunto E de los endomorfismos de G provisto de las dos leyes $(f, g) \mapsto f + g$ y $(f, g) \mapsto f \circ g$. Demostrar que E es un anillo unitario. ¿Posee divisores de cero?
Estudiar el caso de $G = (\mathbf{Z}/2\mathbf{Z}) \times (\mathbf{Z}/2\mathbf{Z})$ (V. cap. 4, ej. 73).

92. Sean A_1 y A_2 dos anillos, se provee $A_1 \times A_2$ de las dos operaciones

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$$
$$(x_1, x_2)(y_1, y_2) = (x_1 y_1, x_2 y_2).$$

a) Demostrar que $A_1 \times A_2$ es un anillo llamado *anillo producto cartesiano* de A_1 y A_2.

b) ¿Cómo son los anillos $A_1 \times A_2$ si A_1 y A_2 son conmutativos? ¿Unitarios? ¿Integros?

93. Sea A un anillo unitario o no, se da al conjunto $A = \mathbf{Z} \times A$ las dos leyes internas

$$(m, a) + (n, b) = (m + n, a + b)$$
$$(m, a)(n, b) = (mn, mb + na + ab).$$

a) *Demostrar que* A′ *es un anillo unitario; ¿cuál es su elemento unidad* e'?

b) Demostrar que $\{0\} \times A$ es un ideal bilateral de A′, isomorfo a A.

c) Demostrar que si A tiene un elemento unidad e, $(0, e)$ es un elemento idempotente (capítulo 3, ej. 51) de A′ y que pertenece al centro de A′.

94. Tenemos un anillo A, se dice que un elemento x de A es *nilpotente* si hay un entero $n > 0$ tal que $x^n = 0$. Demostrar que si x e y son nilpotentes y *permutables*, $x + y$ y xy son nilpotentes (para $x + y$ utilizar la fórmula del binomio).

95. Siendo A un anillo, análogamente lo es el conjunto $A' = \mathfrak{F}(E, A)$ provisto de las dos operaciones $f + g$ y fg (§ 90, ej. 5).

a) Si X es una parte de E, la parte de A′ descrita por las funciones f nulas sobre X es un ideal bilateral de A.

b) En el anillo unitario $A' = \mathfrak{F}(\mathbf{R}, \mathbf{R})$ se considera el ideal $I(x_0)$ descrito por las funciones nulas en x_0 fijado. Dado $x_1 \neq x_2$, caracterizar los ideales intersección y suma de $I(x_1)$ e $I(x_2)$ (V. § 94, ej. 2), demostrar en particular que $I(x_1) + I(x_2) = A$; deducir que $I(x_0)$ es maximal.

Demostrar que $A'/I(x_0)$ es un cuerpo isomorfo a $\mathbf{R}$.

96. Se considera un anillo conmutativo unitario A (elementos neutros 0 y 1) y un elemento fijo d de A. Se representa por $A[\sqrt{d}]$ el conjunto $A \times A$ que posee las dos operaciones

$$(a, a') + (b, b') = (a + b, a' + b')$$
$$(a, a')(b, b') = (ab + da'b', ab' + a'b).$$

a) Demostrar que $A[\sqrt{d}]$ tiene una estructura de anillo conmutativo unitario.

b) Demostrar que A' descrito por $(a, 0)$, cuando a describe A, es un subanillo de $A[\sqrt{d}]$ isomorfo a A. Se *identifica* A' y A (es decir, $(a, 0) = a$).

Demostrar que todo elemento de $A[\sqrt{d}]$ se escribe de una manera única $a + a'\,(0, 1)$.

c) Demostrar que si A es un subanillo de un anillo conmutativo B conteniendo un elemento α tal que $\alpha^2 = d$, $A[\sqrt{d}]$ es el subanillo engendrado por $A \cup \{\alpha\}$, que ha sido obtenido al hacer la *adjunción* de α a A: se dice que es una *extensión cuadrática* de A, se le representa $A[\alpha]$. (Se demostrará primero que $(0, 1)^2 = (d, 0) = d$).

d) Se llama conjugado de $z = x + y\alpha$ ($x, y \in A$) al elemento $\bar{z} = x - y\alpha$ y se designa $N(z) = z\bar{z}$. Demostrar que

$$\overline{z + z'} = \bar{z} + \bar{z}', \qquad \overline{zz'} = \bar{z}\bar{z}', \qquad N(zz') = N(z)N(z').$$

Demostrar que $A[\alpha]$ es íntegro si y sólo si A es íntegro y si para todo z de $A[\alpha]$, $N(z) = 0 \Rightarrow z = 0$. Demostrar que z es inversible en $A[\alpha]$ si y solamente si $N(z)$ es inversible en A. Calcular z^{-1}.

e) Tomando $A = \mathbf{Z}$, $d = -1$ y poniendo $i^2 = -1$ se obtiene el anillo $\mathbf{Z}[i]$ de los *enteros de Gauss.*

Encontrar todos los elementos inversibles de este anillo, verificar que para la multiplicación, describen un grupo de cuatro elementos.

97. Con las mismas notaciones que en el ejercicio 96 se reemplaza el anillo A por un cuerpo conmutativo K.

a) Demostrar que $K[\alpha]$ es un cuerpo conmutativo si y solamente si d no es el cuadrado de un elemento de K ($d = \alpha^2$).

b) Tomando $K = \mathbf{Q}$ y $d = -1$ se obtiene el cuerpo $\mathbf{Q}[i]$, mostrar que $\mathbf{Q}[i]$ es el cuerpo de fracciones del anillo $\mathbf{Z}[i]$.

c) Igualmente si $K = \mathbf{Q}$ y d es un entero natural > 1 no divisible por un cuadrado, $\mathbf{Q}[\sqrt{d}]$ es el cuerpo de las fracciones de $\mathbf{Z}[\sqrt{d}]$.

98. Se dice que un anillo de integridad, unitario, A es *euclídeo* si tiene una aplicación f de A^* en N tal que:

1. Para todo x y todo y de A^*, $f(xy) \geq f(y)$.

2. Para todo a de A y todo b de A* existen dos elementos q y r de A tales que

$$a = bq + r \quad \text{y} \quad (r = 0 \quad \text{o} \quad f(r) < f(b)).$$

el par (q, r) es único.

a) **Demostrar** que si la función f es tal que

$$x \neq y \Rightarrow f(x - y) \leq \sup (f(x), f(y))$$

b) Demostrar que todo anillo euclídeo es principal (considerar un ideal $I \neq \{0\}$ y en I, a tal que $f(a)$ sea mínimo y aplicar la propiedad a x, elemento cualquiera de I, y a a).

c) Demostrar que los anillos siguientes provistos de las aplicaciones f indicadas son euclídeos:

1. $A = \mathbf{Z}$, $f(x) = |x|$ (¿son únicos q y r?).
2. $A = K[X]$, $f(x) =$ grado de x (K[X] es el anillo de los polinomios con coeficientes en el cuerpo conmutativo K; ver capítulo 11).
3. $A = \mathbf{Z}[\sqrt{2}]$, $f(x) = |x\bar{x}|$ (ver ej. 96 y 97).

Se pondrá en $\mathbf{Q}[\sqrt{2}]$, $f(x) = |x\bar{x}|$ y se demostrará que para toda fracción a/b de elementos de $\mathbf{Z}[\sqrt{2}]$ existe un elemento q de $\mathbf{Z}[\sqrt{2}]$ tal que $f(a/b - q) \leq 1/2$. Encontrar q y r para $a = 7 + 3\sqrt{2}$, $b = 2 - \sqrt{2}$.

4. $A = \mathbf{Z}[i]$, $f(x) = x\bar{x}$ (ver ej. 96 y 97).

Se hará en $\mathbf{Q}[i]$, $f(x) = x\bar{x}$ y se demostrará que para toda fracción a/b de elementos de $\mathbf{Z}[i]$ hay un elemento q de $\mathbf{Z}[i]$ tal que $f(a/b - q) \leq 1/2$.
Escribiendo de una manera general $\mathbf{Z} = x + iy$ $(x, y \in \mathbf{Z})$ se podrá también representar z por un punto de un plano, referido a una base ortonormal, llamado *imagen* de z (ver § 117) y determinar el conjunto de las imágenes de xb, iyb y $(x + iy)b$ cuando b es un *elemento fijado* de $\mathbf{Z}[i]$, cuando x e y describen $\mathbf{Z}$. Encontrar todos los pares q y r para $a = 5 + 6i$, $b = 2 + i$.

99*. En $\mathbf{Z}[\sqrt{2}]$ descrito por $z = x + y\sqrt{2}$, se pone $f(z) = |z\bar{z}| = |x^2 - 2y^2|$ (ver ej. 96) y se considera el conjunto U de los elementos inversibles de $\mathbf{Z}[\sqrt{2}]$.

a) Demostrar que z es inversible si y solamente si $f(z) = 1$ (ver ej. 96, **d**). Demostrar que U es un grupo para la multiplicación.

b) Para la relación de orden inducida sobre U por la relación de orden $z \leq z'$ definida en **R**, clasificar $x + \sqrt{2}y$ con relación a 1 y -1 según los signos de x e y.

c) Demostrar que entre los elementos de U estrictamente superiores a 1 hay uno z_0, menor que todos los demás. Determinar z_0.
Demostrar que el conjunto de los elementos positivos de U es un grupo cíclico engendrado por z_0.

d) Resolver en números enteros las ecuaciones

$$1. \quad x^2 - 2y^2 = 1 \qquad\qquad 2. \quad x^2 - 2y^2 = -1.$$

100*. Se llaman *enteros* del cuerpo $\mathbf{Q}[\alpha]$ en que $\alpha^2 = d$, d entero racional no divisible por un cuadrado (V. ej. 96 y 97) los elementos $z = x + y\alpha$ de $\mathbf{Q}[\alpha]$ tales que $z + \bar{z}$ y $z\bar{z}$ son enteros racionales.

a) Demostrar que z es raíz de una ecuación de segundo grado: $x^2 + a_1x + a_2 = 0$ (a_1, a_2 enteros racionales) (es de aquí que proviene la palabra entero; se llama, en

efecto, *número algebraico* toda solución de $a_0x^n + \ldots + a_n = 0$ ($a_0, \ldots, a_n$ *enteros racionales*) y *entero algebraico* toda solución de la ecuación precedente cuando $a_0 = 1$).

b) Demostrar que los enteros de $\mathbf{Q}[\alpha]$ describen un subanillo A de $\mathbf{Q}[\alpha]$, (obsérvese que si z y z' son enteros de $\mathbf{Q}[\alpha]$, entonces $zz' + \bar{z}\bar{z}'$ y $z\bar{z}' + \bar{z}z'$ son enteros racionales).

c) Demostrar que los enteros de $\mathbf{Q}[\alpha]$ describen un grupo aditivo engendrado por:

1. 1 y α si $d \not\equiv 1 \pmod 4$,
2. $1/2\,(1 + \alpha)$ y $1/2\,(1 - \alpha)$ si $d \equiv 1 \pmod 4$)

(se demostrará que si $z = a + b\alpha$ es un entero de $\mathbf{Q}[\alpha]$, $2a$ y $2b$ son enteros racionales que verifican $(2a)^2 - d(2b)^2 \equiv 0$, mod 4, y se determinará su paridad. Se ve, pues, que $A = \mathbf{Z}[\alpha]$ sólo es verdadero en el caso 1.)

d) Demostrar que los únicos elementos inversiibles del anillo A por $d < 0$ son 1 y -1, salvo para $d = -1$ y $d = -3$, encontrarlos todos en estos últimos casos.

101. a y b son enteros naturales primos entre si tales que $b < a$.

a) Demostrar que existen enteros naturales no nulos $a_0, \ldots, a_n$ tales que

$$\frac{a}{b} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots \; a_{n-1} + \cfrac{1}{a_n}}}}.$$

(Utilizar el algoritmo de EUCLIDES para encontrar el m. c. d. de a y de b, § 99, ej. 2.) Se escribirá $a/b = (a_0, a_1, \ldots, a_n)$ (no confundir esta notación con un $(n + 1)$ — étuple o con un ideal), se dice que $(a_0, a_1, \ldots, a_n)$ es una *fracción continua*.

b) Consideremos la fracción continua $(a_0, a_1, \ldots, a_p)$ ($p \leq n$), es igual a una fracción irreducible P_p/Q_p de términos positivos llamada *reducida de rango p de a/b*. Demostrar que

$$P_p = P_{p-1}a_p + P_{p-2}, \qquad Q_p = Q_{p-1}a_p + Q_{p-2}$$
$$P_pQ_{p-1} - P_{p-1}\,Q_p = (-1)^{p-1}$$

¿Cuáles son los valores P_n y Q_n?
¿Qué relación hay entre P_p, Q_p y los enteros U_k, V_k (§ 99, ej. 2, *b*)?

c) Encontrar una pareja de enteros racionales u y v tales que $au + bv = 1$ para $a = 13$, $b = 5$; $a = 75$, $b = 14$.

102 Si a, b, c son enteros racionales, se considera la ecuación en que las incógnitas x e y son enteros racionales

(1) $$ax + by = c.$$

a) Discutir la existencia de las soluciones de (1) (se introducirá el m. c. d. de a y b).

b) Cuando (1) tiene una solución (x_0, y_0) encontrar las demás.

c) Resolver $5x + 13y = 6$ en números enteros (utilizar el ej. 101, c, o el ej. 2 del § 99, o por tanteo con la ayuda del ejercicio 1, *b*, del § 99).

103*. Si p y q son dos enteros primos entre sí:

a) Demostrar que cualesquiera que sean los enteros y, z existe x tal que

$$x \equiv y \pmod{p} \quad \text{y} \quad x \equiv z \pmod{q}$$

(utilizar la igualdad de Bezout; para calcular un x, utilizar el ej. 101 **c**). Ejemplo

$$p = 5, \quad q = 13, \quad y = 3, \quad z = 9.$$

Demostrar que todas las soluciones x son congruentes módulo $n = pq$.

b) Demostrar que a un par $(\dot{y}, \dot{z})$, $\dot{y}$ entero módulo p, $\dot{z}$ entero módulo q, el resultado del apartado **a)** asocia una clase única $\dot{x}$, módulo n. Sea f la aplicación del anillo producto $(\mathbf{Z}/p\mathbf{Z}) \times (\mathbf{Z}/q\mathbf{Z})$ (ver ej. 92) en $\mathbf{Z}/n\mathbf{Z}$, también definida. Demostrar que f es un isomorfismo de anillos (se demostrará primero que es un homomorfismo, después que es inyectiva y se compararán los cardinales de los dos anillos).

104*. Tenemos n que es un entero estrictamente positivo, se designa por $\varphi(n)$, indicador de Euler, el número de los enteros m tales que $1 \leq m \leq n$, y que m y n sean primos entre sí.

a) Demostrar que $\varphi(1) = 1$ y que $\varphi(n)$ es el número de elementos inversibles de $\mathbf{Z}/n\mathbf{Z}$; es también el número de generadores de un grupo cíclico de orden n (ver cap. 4, ej. 74).

b) Demostrar que $n = \Sigma\varphi(d)$, la Σ se extiende a todos los divisores de n.

c) Demostrar que si p y q son enteros positivos primos entre sí.

$$\varphi(pq) = \varphi(p)\varphi(q)$$

(contar los elementos inversibles de los anillos considerados en el ej. 103 **b**).

d) Demostrar que si $n = p^k$ (p primo > 0, $k > 0$)

$$\varphi(n) = \varphi(p^k) = p^{k-1}(p - 1) = n\left(1 - \frac{1}{p}\right)$$

deducir de ello que si $n = p_1^{k_1} \ldots p_m^{k_m}$ es la descomposición de n en factores primos

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right) \ldots \left(1 - \frac{1}{p_m}\right).$$

105. Se considera un anillo *principal* A.

a) Se nos da a_1, a_2, a_n, demostrar que existe D tal que

$$(a_1) + (a_2) + \ldots + (a_n) = (\mathrm{D}).$$

Se dice que D es *un* m.c.d. de $a_1, a_2, \ldots, a_n$, ¿cuáles son los otros? Enunciar y demostrar un teorema análogo al teorema 5′ (§ 100).

b) Se nos dan $a_1, a_2, \ldots, a_n$, demostrar que existe M tal que

$$(a_1) \cap (a_2) \ldots \cap (a_n) = (\mathrm{M}),$$

se dice que M es *un* m.c.m. de $a_1, \ldots, a_n$, ¿cuáles son los demás?

c) Demostrar que en el anillo principal A, las propiedades 1, 2, 3, son equivalentes, igualmente que las propiedades 4, 5, 6 (§ 103, *c*).

106*. **a)** Sea A un anillo conmutativo unitario y una sucesión creciente (para la inclusión) de ideales I_n ($n \in \mathbf{N}$), demostrar que $I = \bigcup_{n \in \mathbf{N}} I_n$ es un ideal de A.

b) Demostrar que en un anillo principal, toda sucesión creciente de ideales es forzosamente estacionaria (ver § 28) (considerar x tal que $(x) = I$, definido en **a**), existe p, tal que $x \in I_p$; observar que $xA = (x)$ está contenido en I_p).

c) Demostrar que toda familia $\mathfrak{J}$ no vacía de ideales de un anillo principal admite un elemento maximal. Deducir que en un anillo principal, todo elemento admite un divisor extremal.

d) Demostrar que todo elemento x de un anillo principal puede escribirse

$$x = p_1^{k_1} \dots p_n^{k_n}$$

de donde $p_1, \dots, p_n$ son elementos extremales dos a dos no asociados y $k_1, \dots, k_n$ enteros estrictamente positivos, y esto de una manera única a condición de poder sustituir cualquiera p_i por un elemento asociado.
Un anillo unitario íntegro que tiene esta última propiedad se llama anillo *factorial*; hay anillos factoriales que no son principales (V. cap. 11, ej. 355).

107. Se considera el anillo $A = \mathbf{Z}[\alpha]$ donde $\alpha^2 = -5$ y se escribe $z = x + y\alpha$; $N(z) = z\bar{z}$ (V. ej. 96).

a) Demostrar que los elementos inversibles de A son $+1$ y -1.

b) Demostrar que 3, $2 + \alpha$, $2 - \alpha$ son elementos extremales de A.

c) De la igualdad $3 \cdot 3 = (2 + \alpha)(2 - \alpha)$, deducir que A no es un anillo principal.

108. Dado un anillo conmutativo A y un ideal I de A, se designa por f el homomorfismo canónico de A sobre $A' = A/I$. Si J' es un ideal de A' se escribe $J = f^{-1}(J')$.

a) Demostrar que J es un ideal de A conteniendo I.

b) Demostrar que $J' = f(J)$; deducir que la aplicación del conjunto de los ideales de A que contienen a I, en el conjunto de los ideales de A' definida por $J \to f(J)$ es una biyección, estrictamente creciente (para la inclusión de los conjuntos).

c) Demostrar que A/J y A'/J' son isomorfos.

109*. Sea A un anillo conmutativo unitario e I un ideal de A.

a) Demostrar que A/I es un cuerpo si y solamente si I es maximal (utilizar el ejercicio 108 **b** y el ejercicio del § 106).

b) Demostrar que todo ideal maximal es primo (utilizar § 99, ej. 4. La recíproca es en general falsa, pero en un anillo principal, todo ideal *propio* primo es maximal, ver ej. 105, **c**); por ejemplo, en **Z**, (0) es primo maximal).

110. Se considera el anillo $A = 2\mathbf{Z}$.

a) Demostrar que A es íntegro y que todo ideal A es principal.

b) Determinar los ideales maximales de A; demostrar que hay uno y sólo uno tal que A/J no sea un cuerpo (A/J posee incluso divisores de cero) (A no es unitario, el resultado del ej. 109 **a**) no es aplicable).

111. Se considera el conjunto S de las sucesiones *racionales* (u_n) provisto de las dos operaciones

$$(u_n) + (v_n) = (u_n + v_n), \qquad (u_n)(v_n) = (u_n v_n)$$

S es, pues, $\mathfrak{F}(\mathbf{N}, \mathbf{Q})$ (ver ej. 95). Se consideran los subconjuntos de S siguientes:

B conjunto de las sucesiones acotadas.
A — — de Cauchy (Agustín).
C — — convergentes.
Z — — convergentes hacia cero.
N — — tales que u_n es nulo a partir de un cierto orden.

a) Demostrar que N, Z, C, A, B, S es una sucesión de anillos *estrictamente* creciente, cada uno de ellos, salvo S, es un subanillo de los siguientes. ¿Cuáles son los anillos unitarios? ¿Tienen estos anillos divisores de cero?

b) Demostrar que N (resp. Z) es un ideal de B (luego anillos comprendidos entre N (resp. Z) y B).
Demostrar que N es un ideal de S, y, en consecuencia, de todos los demás.

c) Demostrar que Z es un ideal maximal de A (si I es un ideal de A tal que Z esté estrictamente incluido en I, se observará que I contiene una sucesión (a_n) que no tiende a cero y se demostrará que I es igual a A). ¿Cuál es el anillo cociente A/Z?

d) Demostrar que Z es un ideal maximal y primo de C (se demostrará que es maximal como en **c)** y se utilizará el ej. 109, **b)**).
Demostrar que Z es un ideal no primo y no maximal de B (considerar las sucesiones definidas por $a_n = 1 + (-1)^n$, $b_n = 1 - (-1)^n$).
(Este ejercicio muestra que las propiedades «I es un ideal», «I es un ideal primo», «I es un ideal maximal» son verdaderas y falsas para $I \subset A$ según el anillo A que analizamos; ver § 94, observación.)

112. *Buscar las soluciones enteras de las congruencias siguientes (cuando son resolubles):*

a) $5x \equiv 12 \pmod 7$.
b) $4x \equiv 9 \pmod 8$.
c) $1963x \equiv 2000 \pmod{1964}$,

si a, b, n son enteros dados, discutir la existencia de las soluciones enteras de $ax \equiv b \pmod n$.

113. **a)** Sabiendo que los coeficientes y las incógnitas pertenecen a un anillo conmutativo unitario A, en qué condición el sistema siguiente tiene solución

$$ax + by = c \qquad a'x + b'y = c'.$$

Caso particular: A es un cuerpo conmutativo.

b) Buscar las soluciones enteras de los sistemas siguientes cuando tienen solución

1. $7x + 5y \equiv 2 \pmod 8$ $\quad 5x + 4y \equiv 16 \pmod 8$.
2. $7x + 5y \equiv 2 \pmod 9$ $\quad 5x + 4y \equiv 16 \pmod 9$.

Buscar el entero λ para que el sistema

$$7x + 5y \equiv 2 \pmod 8 \qquad 5x + \lambda y \equiv 16 \pmod 8$$

tenga solución.

114. Llamamos sección inicialmente abierta s, toda parte de $\mathbf{Q}$ poseyendo las propiedades siguientes:

1. *s es una parte propia* ($s \neq \varnothing$, $s \neq \mathbf{Q}$).
2. $x \in s$ y $y \leq x \Rightarrow y \in s$.
3. s no tiene elemento máximo.

Se designa por S el conjunto de las secciones que empiezan abiertas de **Q**.

a) Demostrar que el conjunto de los racionales x estrictamente inferior a un racional a es una sección inicialmente abierta, que se la designará $s(a)$.

b) Demostrar que la relación $s_1 \subset s_2$ es *una relación de orden total definida sobre* S, se la representará $s_1 \leq s_2$.

c) Si (s_i) es una familia cuyo conjunto de índices es I, de elementos de S acotados superiormente por s_0, determinar que $s = \bigcup_{i \in I} s_i$ es una sección inicialmente abierta y que es la menor que contiene cada s_i. De ello deducir que en S totalmente ordenado por $s_1 \leq s_2$, toda parte acotada superiormente tiene un límite superior.

d) Demostrar que dado s_1 y s_2, $s_1 < s_2$, hay un racional x tal que $s_1 < s(x) < s_2$.

115. Sean s_1 y s_2 (ver ej. 114) *descritas, respectivamente,* por los racionales x_1 y x_2; se designa $s_1 + s_2$ la parte de **Q** descrita por $x_1 + x_2$.

a) Demostrar que $s_1 + s_2 \in$ S se define también una adición en S; demostrar que es conmutativa y asociativa.

Demostrar que $s(x_1) + s(x_2) = s(x_1 + x_2)$.

b) Demostrar que $s(0)$ es el elemento neutro de esta adición. Dado s de **S**, demostrar que el conjunto de los racionales x tales que $s < s(-x)$ es una sección inicialmente abierta s' verificando $s + s' = s(0)$, se la representa $-s$. (Utilizar el hecho que el orden definido sobre **Q** es arquimediano.)

c) Demostrar que para todo s de S

$$s_1 \leq s_2 \Rightarrow s_1 + s \leq s_2 + s.$$

116. Los resultados de los ejercicios 114 y 115 dan, pues, a S una estructura de grupo aditivo totalmente ordenado.

a) Demostrar que S′ descrito por las secciones inicialmente abiertas $s(x)$, x describiendo **Q** es isomorfo para la adición y el orden a **Q**, se identificará x y $s(x)$, en consecuencia $\mathbf{Q} \subset$ S.

Se llamará *número real* cada elemento de S (su conjunto se designa por **R**), es o bien *racional* (si pertenece a **Q**), o bien *irracional*, si pertenece al complementario de **Q** con relación a **R**.

b) Demostrar que el orden sobre **R** ($s_1 \leq s_2$) es arquimediano, que entre dos reales cualesquiera, hay un racional y que entre dos racionales, hay un irracional.

117. Si s es un real estrictamente positivo ($s > 0$), se escribe (ver ej. 114 a 116)

$$s' = s \cap \mathbf{Q}^*_+.$$

a) $s'_1 s'_2$ designa el conjunto descrito por $x_1 x_2$, donde x_1 y x_2 describen, respectivamente, s'_1 y s'_2, demostrar que $s'_1 s'_2 \cup \mathbf{Q}$ es un real > 0 (es decir, una sección inicialmente abierta) que se designará por $s_1 s_2$, se define así una multiplicación en $\mathbf{R}^*$; demostrar que es conmutativa, asociativa y distributiva respecto a la suma y que si x_1 y x_2 son racionales estrictamente positivos

$$s(x_1)s(x_2) = s(x_1 x_2).$$

b) *Demostrar que para todo s de R_+^* $ss(1) = s$.*

c) Si x describe $s' = s \cap \mathbf{Q}_+^*$ demostrar que el complemento respecto a $\mathbf{Q}_+^*$ del conjunto descrito por x^{-1} es una sección inicialmente abierta s'' verificando que $ss'' = s(1)$, se le representará s^{-1}.

d) Demostrar que $s_1 > 0$ y $s_2 > 0$ implican $s_1 s_2 > 0$.

e) Extender a $\mathbf{R}$ la multiplicación definida en $\mathbf{R}_+^*$ (V. § 62, la prolongación de $\mathbf{Z}$ de la multiplicación en $\mathbf{N}$) imponiéndole la condición de ser distributiva respecto a la suma; demostrar la regla de los signos; demostrar que para todo s de $\mathbf{R}$ $ss(0) = s(0)$.

118. Demostrar que todo real $s > 0$ (ver ej. 114 al 117) admite una *raíz n-ésima única* estrictamente positiva δ (examinar el conjunto de los racionales positivos x tales que $s(x^n) < s$ y utilizar el hecho que toda parte acotada superiormente de $\mathbf{R}$ tiene un límite superior). Discutir la existencia y el número de las raíces n-ésimas de un número real cualquiera.

119*. Demostrar que el conjunto E de los axiomas de $\mathbf{R}$, E_1, E_2, E_3, E_4 (§ 109) es equivalente al conjunto E′ de los axiomas E_1, E_2, E_3, E_4'. E_4': «la intersección de una sucesión $[a_n, b_n]$ ($n \in \mathbf{N}$) de intervalos cerrados que verifican para todo n, $a_n \leq a_{n+1}$ y $b_{n+1} \leq b_n$ es no vacío».
(Para $E \Rightarrow E'$ introducir sup a_n e inf b_n. Para demostrar que $E' \Rightarrow E$ considerar una parte X *no vacía acotada superiormente* de $\mathbf{R}$ y el conjunto M de sus cotas superiores. Si $a \in X$, hay un entero natural mínimo p_n tal que $a + p_n 2^{-n} \in M$. Considerar los intervalos $I_n = [a + (p_n - 1)2^{-n},\ a + p_n 2^{-n}]$ y su intersección J (no vacía) y demostrar que J no puede contener dos elementos distintos.)

120. Sea G un subgrupo aditivo de $\mathbf{R}$, se designa por G′ el conjunto de los elementos *estrictamente* positivos de G y se pone $m = \inf G'$.

a) Demostrar que $m \geq 0$.

b) Demostrar que G pertenece a uno de los dos tipos siguientes
1. $m \in G'$ entonces $m > 0$ y $G = m\mathbf{Z}$.
2. $m \notin G'$, entonces $m = 0$ y para todo x real y todo h real estrictamente positivo hay al menos un elemento de G en $]x - h, x + h[$.

c) Si (u, v) describen $\mathbf{Z} \times \mathbf{Z}$ y a y b son dos números reales *dados*, demostrar que el conjunto descrito por $ua + vb$ es un subgrupo G de $\mathbf{R}$.
¿Qué propiedad de (a, b) permite decidir si G es del tipo 1 o del tipo 2?

121. Si x es un número real, se designa por $[x]$, llamada *parte entera* de x, el mayor entero racional contenido en x, es decir, el entero p tal que $p \leq x < p + 1$. Si x e y son reales y n un entero estrictamente positivo, demostrar los resultados siguientes

a) $[x + y] = [x] + [y] + \varepsilon$ con $\varepsilon = 0$ o $\varepsilon = 1$.

b) $[x - y] = [x] - [y] - \varepsilon$ con $\varepsilon = 0$ o $\varepsilon = 1$.

c) $[x] + \left[x + \dfrac{1}{n}\right] + \dots + \left[x + \dfrac{n-1}{n}\right] = [nx]$.

d) $\left[\dfrac{[nx]}{n}\right] = [x]$.

6

NUMEROS COMPLEJOS

I. El cuerpo de los números complejos. Módulo de un número complejo.
II. Representación geométrica de un número complejo. Argumento de un número complejo.
III. Aplicaciones de los números complejos.

I. El cuerpo de los números complejos. Módulo de un número complejo

112. Introducción

Hemos visto que en **R** los números estrictamente negativos no tienen raíz cuadrada. Nos proponemos determinar un supercuerpo conmutativo de **R**, K, en el que todo elemento admita una raíz cuadrada. En particular, -1 deberá tener una raíz cuadrada, sea i una de ellas; luego $i^2 + 1 = 0$.

Vamos a demostrar que si existe K, el subcuerpo K′ engendrado por $\mathbf{R} \cup \{i\}$ está descrito por los elementos de la forma $a + a'i$, en donde a y a' describen **R**. Desde luego

(1) $$a + a'i = 0 \Rightarrow a = a' = 0$$

si no para $a' \neq 0$, i raíz cuadrada de -1 pertenecería a **R**. Por otro lado, las reglas de cálculo en un cuerpo conmutativo prueban que, teniendo en cuenta $i^2 + 1 = 0$,

(2) $$(a + a'i) + (b + b'i) = (a + b) + (a' + b')i$$

(3) $$(a + a'i)(b + b'i) = (ab - a'b') + i(ab' + a'b).$$

Las igualdades (1) y (2) demuestran que

(4) $$(a + a'i = b + b'i) \Leftrightarrow (a = b,\ a' = b').$$

En fin, la igualdad

$$(a + a'i)(a - a'i) = a^2 + a'^2$$

lo que demuestra que, si $a + a'i \neq 0$, es

$$(a + a'i)^{-1} = \frac{a - a'i}{a^2 + a'^2} = \frac{a}{a^2 + a'^2} + \frac{-a'}{a^2 + a'^2} i$$

pues en **R**: $a^2 + a'^2 = 0$, equivale a $a = a' = 0$; luego $a + a'i = 0$.

Luego si K existe, K′ descrito por $a + a'i$ es un cuerpo. Observemos que i debe cumplir las condiciones: i pertenece a K e $i^2 + 1 = 0$, si existe otro elemento i_1 de K tal que $(i_j)^2 + 1 = 0$, se ve que el subcuerpo K'_1 de K, engendrado por $\mathbf{R} \cup \{i_1\}$ es isomorfo a K′: no podemos determinar K′ más que salvo un isomorfismo. Vamos a construir un cuerpo, el cuerpo **C** de los números complejos isomorfo a K′; demostraremos seguidamente que el cuerpo obtenido responde a la pregunta: todo elemento y admite una raíz cuadrada (§ 114); veremos incluso en el § 119 que todo elemento de **C** admite raíces n-ésimas en **C** y en el capítulo 11 que toda ecuación con coeficientes $a_0, \ldots, a_n$ en **C** de la forma

$$a_0 x^n + a_1 x^{n-1} + \ldots + a_n = 0$$

admite al menos una solución en el cuerpo **C**.

La igualdad (4) muestra que hay una biyección entre el conjunto K′ —si existe— y $\mathbf{R} \times \mathbf{R}$; además, la igualdad (2) demuestra que existe un *isomorfismo entre el grupo aditivo K′ y el grupo aditivo* $\mathbf{R} \times \mathbf{R}$ *provisto de la adición* (ver § 79)

(2′) $$(a, a') + (b, b') = (a + b, a' + b');$$

en efecto,

$$\left\{\begin{array}{l} a + a'i \to (a, a') \\ b + b'i \to (b, b') \end{array}\right. \Rightarrow (a + a'i) + (b + b'i) \to (a, a') + (b, b')$$

guiados por (3) proveemos, además, $\mathbf{R} \times \mathbf{R}$ de la multiplicación

(3′) $$(a, a')(b, b') = (ab - a'b', ab' + a'b).$$

La igualdad (3) demuestra entonces

$$(a + a'i)(b + b'i) \to (a, a')(b, b').$$

Luego si demostramos que **C**, es decir, el conjunto $\mathbf{R} \times \mathbf{R}$ provisto de las operaciones (2′) y (3′) es un cuerpo, habremos probado la existencia de K′, esto es lo que vamos a hacer.

113. El cuerpo de los números complejos

Sea **C** el conjunto $\mathbf{R} \times \mathbf{R}$ provisto de dos operaciones definidas por las igualdades (2′) y (3′) del parágrafo precedente.

La suma (2′) da a este conjunto una estructura de *grupo aditivo*, el elemento *cero* es (0, 0) y el *opuesto* de (a, a'), es el elemento $(-a, -a')$.

La multiplicación (3′) es *asociativa*

$$\begin{aligned}\lambda &= [(a, a')(b, b')](c, c') = (ab - a'b', ab' + a'b)(c, c')\\ &= [(ab - a'b')c - (ab' + a'b)c', (ab - a'b')c' + (ab' + a'b)c]\\ \mu &= (a, a')[(b, b')(c, c')] = (a, a')[bc - b'c', bc' + b'c]\\ &= [a(bc - b'c') - a'(bc' + b'c), a(bc' + b'c) + a'(bc - b'c')]\end{aligned}$$

se ve fácilmente que $\lambda = \mu$.

La fórmula (3′) muestra inmediatamente que la multiplicación es *conmutativa*; finalmente es *distributiva* respecto a la suma; en efecto,

$$\begin{aligned}&(a, a')[(b, b') + (c, c')] = (a, a')[b + c, b' + c']\\ &= [a(b + c) - a'(b' + c'), a(b' + c') + a'(b + c)]\\ &= (ab - a'b', ab' + a'b) + (ac - a'c', ac' + a'c) = (a, a')(b, b') + (a, a')(c, c').\end{aligned}$$

Busquemos si existe un elemento *unidad* (x, x'); para todo (a, a') tendremos

$$(a, a')(x, x') = (a, a') \Leftrightarrow \begin{cases} ax - a'x' = a\\ ax' + a'x = a'.\end{cases}$$

De donde $(x, x') = (1, 0)$.

En fin, para todo $(a, a') \neq (0, 0)$, hay un inverso (b, b'), pues

$$(a, a')(b, b') = (1, 0) \Leftrightarrow \begin{cases} ab - a'b' = 1\\ ab' + a'b = 0.\end{cases}$$

De donde

$$b = \frac{a}{a^2 + a'^2}, \qquad b' = \frac{-a'}{a^2 + a'^2},$$

es decir,

$$(a, a')^{-1} = \left(\frac{a}{a^2 + a'^2}, \qquad \frac{-a'}{a^2 + a'^2}\right)$$

pues (§ 109, *b*) en el cuerpo de los reales $(a, a') \neq 0$ es equivalente a $a^2 + a'^2 \neq 0$.

Luego el conjunto $\mathbf{R} \times \mathbf{R}$ provisto de las operaciones (2′) y (3′) es un *cuerpo conmutativo* **C**.

Veamos finalmente que la parte de **C** descrita por $(a, 0)$ es un subcuerpo de **C** isomorfo a **R**. En efecto,

$$\begin{aligned}a \to (a, 0), \qquad & b \to (b, 0)\\ a + b \to (a + b, 0) &= (a, 0) + (b, 0)\\ ab \to (ab, 0) &= (a, 0)(b, 0)\end{aligned}$$

luego:

TEOREMA Y DEFINICIÓN. — *El conjunto* $\mathbf{R} \times \mathbf{R}$ *provisto de las dos operaciones*

$$(a, a') + (b, b') = (a + b, a' + b')$$
$$(a, a')(b, b') = (ab - a'b', ab' + a'b)$$

es un cuerpo conmutativo cuyo subcuerpo descrito por $(a, 0)$ *es isomorfo a* $\mathbf{R}$. *Sus elementos neutros son* $(0, 0)$ *y* $(1, 0)$. *Se le llama el cuerpo* $\mathbf{C}$ *de los* números complejos[19].

Identifiquemos a de $\mathbf{R}$ y $(a, 0)$ de $\mathbf{C}$; luego $(0, 0) = 0$ y $(1, 0) = 1$; $\mathbf{R}$ es entonces una parte de $\mathbf{C}$. Observemos que para b real

$$b(a, a') = (b, 0)(a, a') = (ba, ba').$$

En consecuencia,

$$(a, a') = (a, 0) + (0, a') = a + a'(0, 1) \tag{5}$$

esta descomposición es *única* (ver § 81). Por otro lado,

$$(0, 1)(0, 1) = (-1, 0) = -1$$

tradicionalmente se escribe $(0, 1) = i$; *todo número complejo* α *se escribe, pues, de una manera única bajo la forma*

$$\alpha = (a, a') = a + a'i \qquad i^2 = -1.$$

Volvemos a encontrar así el subcuerpo K′ (del cuerpo K del que supusimos la existencia).

OBSERVACION *sobre los órdenes eventuales definidos sobre el conjunto* $\mathbf{C}$.

El conjunto $\mathbf{C}$ equipotente a $\mathbf{R} \times \mathbf{R}$ puede ser ordenado de varias maneras. Se ve fácilmente, por ejemplo, que las relaciones de orden (§ 24 y § 108, ej. 4)

$$(R_1) \quad (a, a') \leq (b, b') \Leftrightarrow (a \leq b, a' \leq b')$$
$$(R_2) \quad (a, a') \leq (b, b') \Leftrightarrow (a < b \quad \text{o} \quad a = b, a' \leq b')$$

proveen al grupo aditivo $\mathbf{C}$ de una estructura de *grupo ordenado* (parcial para R_1, total para R_2).

Por el contrario, no existe relación de orden total sobre $\mathbf{C}$ tal que $\mathbf{C}$ sea un *cuerpo ordenado*: en efecto, $1 = 1^2$ y $-1 = i^2$ son cuadrados en $\mathbf{C}$, deberían ser entonces los dos estrictamente superiores a cero, no siendo nulos; ahora bien, -1 y $+1$ son opuestos; es, pues, imposible que sean los dos estrictamente superiores a cero (§ 108).

114. Raíces cuadradas de un número complejo. Ecuación de segundo grado

Hemos construido el cuerpo $\mathbf{C}$ como supercuerpo de $\mathbf{R}$ en el que -1 tenga al menos una raíz cuadrada. Vamos a demostrar que todo número complejo tiene dos raíces cuadradas y demostraremos en el § 119 que todo número complejo tiene n raíces n-ésimas.

(19) Se dice también de los «números imaginarios».

a) Sea el número complejo $a + ib$, demostremos que tiene dos raíces opuestas; sea $x + iy$ una de estas raíces

$$(x + iy)^2 = a + ib \Leftrightarrow x^2 - y^2 + 2ixy = a + ib$$

tenemos que resolver el sistema de coeficientes e incógnitas reales

$$\begin{cases} x^2 - y^2 = a \\ \\ 2xy = b \end{cases} \Leftrightarrow \begin{cases} x^2 + (-y^2) = a \\ x^2(-y^2) = -\dfrac{b^2}{4} \\ xyb \geq 0 \end{cases}$$

x^2 y $(-y^2)$ son, en consecuencia, raíces de la ecuación de coeficientes e incógnitas reales

$$u^2 - au - \frac{b^2}{4} = 0;$$

esta ecuación tiene una raíz positiva x^2 y una raíz negativa $(-y^2)$; entonces, sin ambigüedad,

$$x^2 = \frac{\sqrt{a^2 + b^2} + a}{2} \qquad y^2 = \frac{\sqrt{a^2 + b^2} - a}{2}$$

$$x = \varepsilon \sqrt{\frac{\sqrt{a^2 + b^2} + a}{2}} \qquad y = \varepsilon' \sqrt{\frac{\sqrt{a^2 + b^2} - a}{2}}$$

con $\varepsilon^2 = \varepsilon'^2 = 1$, siendo ε y ε' tales que $xyb > 0$, luego tales que $\varepsilon\varepsilon' b > 0$; de donde obtenemos dos soluciones

$$z_1 = \varepsilon_1 \sqrt{\frac{\sqrt{a^2 + b^2} + a}{2}} + i\varepsilon'_1 \sqrt{\frac{\sqrt{a^2 + b^2} - a}{2}}$$

una de ellas (con $\varepsilon_1\varepsilon'_1$ del signo de b), y $z_2 = -z_1$ la otra.

Para $b = 0$ y $a > 0$ se encuentra

$$z_1 = \varepsilon_1 \sqrt{a} \qquad z_2 = -z_1$$

para $b = 0$ y $a < 0$ (atención: $\sqrt{a^2} = |a| = -a$) se encuentra

$$z_1 = \varepsilon'_1 i \sqrt{-a} \qquad z_2 = -z_1 .$$

Si $a = b = 0$ se encuentra $z_1 = z_2 = 0$, luego:

TEOREMA. — *Todo número complejo tiene dos raíces cuadradas opuestas. Son distintas entre sí, si y sólo si el número es no nulo.*

b) Consideremos una *ecuación de segundo grado con coeficientes complejos* ($a \neq 0$)

$$ax^2 + bx + c = 0$$

se puede escribir

$$\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2} = 0.$$

Se presentan dos tipos de discusión según sean a, b, c reales o no:

(1) *a, b, c reales*

$$b^2 - 4ac > 0 \qquad x = \frac{-b + \varepsilon\sqrt{b^2 - 4ac}}{2a} \qquad (\varepsilon^2 = 1)$$

$$b^2 - 4ac < 0 \qquad x = \frac{-b + \varepsilon i\sqrt{4ac - b^2}}{2a} \qquad (\varepsilon^2 = 1)$$

$$b^2 - 4ac = 0 \qquad x = -\frac{b}{2a}.$$

(2) *a, b, c no reales.* Sea d un número complejo tal que

$$d^2 = b^2 - 4ac$$

$$b^2 - 4ac \neq 0 \qquad x = \frac{-b + \varepsilon d}{2a} \qquad (\varepsilon^2 = 1)$$

$$b^2 - 4ac = 0 \qquad x = -\frac{b}{2a}.$$

EJERCICIOS

1. Encontrar las raíces cuadradas de $5 - 6i$, $4ab + 2(a^2 - b^2)i$ (a y b reales).
2. Resolver las ecuaciones

$$x^2 - (5 - 14i)x - 2(5i + 12) = 0, \qquad x^2 - 2(1 + ia^2)x + 1 - a^4 = 0 \qquad (a \text{ real}).$$

3. Siendo a real y b complejo, resolver $x^2 - 2abx + b^2 = 0$.

115. Números complejos conjugados. Reglas de cálculo

a) En $z = x + iy$ (x e y reales), x se llama *parte real* de z e y *parte imaginaria* de z (esta última denominación es un abuso de lenguaje, pues y es *real*); se dice también que y es el coeficiente de i.

Se utiliza las notaciones siguientes

$$z = x + iy \Rightarrow (x = \Re z,\ y = \Im z).$$

Un número complejo de la forma iy (y real) se llama *complejo puro*, o imaginario puro.

b) Se llama número *complejo conjugado* de $z = x + iy$ el número complejo $x - iy$ que se representa $\bar{z}$. Tenemos

$$\mathcal{R}(\bar{z}) = \mathcal{R}(z), \qquad \mathcal{J}(\bar{z}) = -\mathcal{J}(z)$$

$$\mathcal{R}(z) = \frac{1}{2}(z + \bar{z}), \qquad \mathcal{J}(z) = \frac{1}{2i}(z - \bar{z}).$$

Un número complejo es *real* si y sólo si $z = \bar{z}$.

Un número complejo es *complejo puro* si y sólo si $z + \bar{z} = 0$.

Por otra parte, para todo z y todo z' de **C**

$$\overline{(\bar{z})} = z, \qquad \overline{(z + z')} = \bar{z} + \bar{z}', \qquad \overline{(zz')} = \bar{z}\bar{z}'$$

luego la aplicación $z \to \bar{z}$ es un *automorfismo* del cuerpo **C** *que deja invariante cada número real*. Además, es involutivo (§ 15, ej. 6).

EJERCICIO

Demostrar que todo endomorfismo f de un cuerpo **C** que deja **R** invariante es una de las aplicaciones siguientes: para todo z de **C**

a) $f(z) = 0$, *b*) $f(z) = z$, *c*) $f(z) = \bar{z}$

(utilizar el ejercicio del § 109, *b*).

c) Las fórmulas (2), (3) y (2') y (3') del § 112 muestran que las sumas, sustracciones y multiplicaciones se efectúan en **C** siguen *las reglas de cálculo de un cuerpo conmutativo reemplazando i^2 por* -1. Para la división observemos que

$$z\bar{z} = (x + iy)(x - iy) = x^2 + y^2$$

luego si z es un número complejo *no nulo* y z' un número complejo cualquiera

$$\frac{1}{z} = z^{-1} = \frac{z}{x^2 + y^2}; \qquad \frac{z'}{z} = z'z^{-1} = \frac{z'\bar{z}}{x^2 + y^2}.$$

EJERCICIO

Si z y z' son complejos, se considera los tres números

$$X = \frac{z + z'}{1 + zz'}, \qquad Y = i\,\frac{z' - z}{1 + zz'}, \qquad Z = \frac{1 - zz'}{1 + zz'}.$$

Demostrar que $X^2 + Y^2 + Z^2 = 1$. Demostrar que X, Y, Z son reales si y sólo si $z' = \bar{z}$.

116. Módulo o valor absoluto de un número complejo

a) Consideremos la aplicación de **C** en $\mathbf{R}_+$ definida por

$$z = x + iy \to |z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}$$

$|z|$ es el módulo de z.

Observemos que

$$|z| = |-z| = |\bar{z}|, \qquad |x| \leq |z|, \qquad |y| \leq |z|.$$

Esta aplicación posee las propiedades siguientes

(1) $$z = 0 \Leftrightarrow |z| = 0$$

(2) $$|zz'| = |z|\,|z'|$$

(3) $$|z + z'| \leq |z| + |z'|$$

(1) es evidente, ya que en el cuerpo de los reales $x^2 + y^2 = 0$ es equivalente a $x = y = 0$. Para (2) y (3) es suficiente comparar los cuadrados de dos miembros (siendo éstos positivos)

$$|zz'|^2 = (zz')(\overline{zz'}) = (zz')(\bar{z}\bar{z}') = (z\bar{z})(z'\bar{z}') = |z|^2\,|z'|^2$$

$$|z + z'|^2 = (z + z')\overline{(z + z')} = (z + z')(\bar{z} + \bar{z}') = z\bar{z} + z\bar{z}' + \bar{z}z' + z'\bar{z}'$$
$$= |z|^2 + 2\mathcal{R}(z\bar{z}') + |z'|^2$$

pues el conjugado de $z\bar{z}'$ es $\bar{z}z'$; según las observaciones anteriores

$$\mathcal{R}(z\bar{z}') \leq |z\bar{z}'| = |z|\,|z'|$$

luego

$$|z + z'|^2 \leq |z|^2 + |z'|^2 + 2|z|\,|z'| = (|z| + |z'|)^2.$$

De (3) se deduce fácilmente que

$$|z_1 + z_2 + \ldots + z_n| \leq |z_1| + |z_2| + \ldots + |z_n|.$$

Se ve luego que la aplicación $z \to |z|$ de **C** en $\mathbf{R}_+$ es un valor absoluto (§ 110, definición 2); por otra parte, la aplicación $(z, z') \to |z - z'|$ de $\mathbf{C} \times \mathbf{C}$ en $\mathbf{R}_+$ es una distancia (§ 110, definición 3):

TEOREMA. — *Provisto de la aplicación* $z \to |z| = \sqrt{z\bar{z}}$, **C** *es un cuerpo valorado. Provisto de la distancia* $d(z, z') = |z - z'|$, **C** *es un espacio métrico.*

Se puede llamar a $|z|$ *valor absoluto* de z, la costumbre hace que se prefiera a menudo el término *módulo.*

b) Para todo número complejo z no nulo y todo número complejo z'

$$z^{-1} = \frac{z}{z\bar{z}} = \frac{\bar{z}}{|z|^2}, \qquad \frac{z'}{z} = z'z^{-1} = \frac{z'\bar{z}}{|z|^2}.$$

Tenemos en particular el resultado siguiente:

Un número complejo u tiene por módulo 1 si y sólo si $\bar{u} = u^{-1}$; además, cualquiera que sean u y v de módulo 1

$$|vu^{-1}| = |v|\,|u^{-1}| = |v|\,|\bar{u}| = 1.$$

TEOREMA. — *El conjunto de los números complejos de módulo* 1 *provisto de la multiplicación es un grupo, se le designa por* **U**; *es un subgrupo del grupo multiplicativo* $\mathbf{C}^*$.

En particular,

$$i^2 = -1 \qquad i^3 = -i \qquad i^4 = 1.$$

De una manera más general para todo entero n

$$i^{4n} = 1 \qquad i^{4n+1} = i \qquad i^{4n+2} = -1 \qquad i^{4n+3} = -i.$$

Luego i engendra un subgrupo multiplicativo de orden 4 *de* **U**.

Sea z un número complejo no nulo, el módulo de $z\,|z|^{-1}$ es 1, luego *todo número complejo no nulo puede escribirse de una manera única $z = |z|\,u$, siendo u un número complejo de módulo* 1.

EJERCICIOS

1. Demostrar que $\big|\,|z| - |z'|\,\big| \leq |z + z'|$
2. Demostrar que todo número complejo de módulo 1 puede escribirse de una manera única bajo la forma $(1 - ia)/(1 + ia)$, $(a \in \mathbf{R})$.

II. Representación geométrica de un número complejo. Argumento de un número complejo

En esta sección utilizaremos nociones bien conocidas del lector: *punto, vector ligado, vector libre, traslación, homotecia* en el plano afín de dos dimensiones; a continuación: *longitud* de un vector, *ángulo* de dos vectores, *rotación* en el plano euclídeo orientado de dos dimensiones. Trataremos estas cuestiones de una manera más general en este libro (capítulo 15, § 235) y en el tomo III (Geometría) cuando estudiemos los espacios reales de dimensión n.

Naturalmente, este estudio será independiente de los resultados que vamos a obtener aquí para los números complejos.

117. Interpretación de los números complejos mediante vectores o puntos del plano

Llamamos *punto* del plano $\mathbf{R} \times \mathbf{R}$ toda pareja (x, y). Así, O, A y B son, respectivamente, los puntos (0, 0), (1, 0) y (0, 1). A todo vector libre $\vec{V}$, o a

todo punto M se le puede hacer corresponder (x, y) o $z = x + iy$ biunívocamente mediante las fórmulas

$$\vec{V} = x\overrightarrow{OA} + y\overrightarrow{OB}, \qquad \overrightarrow{OM} = x\overrightarrow{OA} + y\overrightarrow{OB}$$

x e y son las *coordenadas* (*x abscisa, y ordenada*) del vector $\vec{V}$ o del punto M.

Notaremos esta biyección

$$z \to \vec{V} = \overrightarrow{OM}.$$

Se dice que $z = x + iy$ es el *afijo* del vector $\vec{V}$ o del punto M y que V y M son, respectivamente, el *vector imagen* y el *punto imagen* del número complejo z (fig. 11), se designará a este vector $\vec{V}(z)$ y a este punto **M(z).**

Los números reales tienen por imágenes los puntos de Ox que se le llama el *eje real* y los números complejos puros mediante los puntos del eje Oy que se llama el *eje imaginario*; el plano **R** × **R** imagen del conjunto **C** de los números complejos se llama el *plano complejo.* Estas dos denominaciones: eje imaginario y plano complejo son *abusos de lenguaje,* tanto el eje como el plano están descritos por puntos de coordenadas reales.

Las definiciones de la adición de los vectores libres y de la suma de los números complejos muestran que el grupo aditivo de los vectores libres es isomorfo al grupo aditivo **C** (fig. 12)

(1)
$$\begin{cases} z \to \vec{V} \\ z' \to \vec{V}' \end{cases} \Rightarrow (z + z' \to \vec{V} + \vec{V}').$$

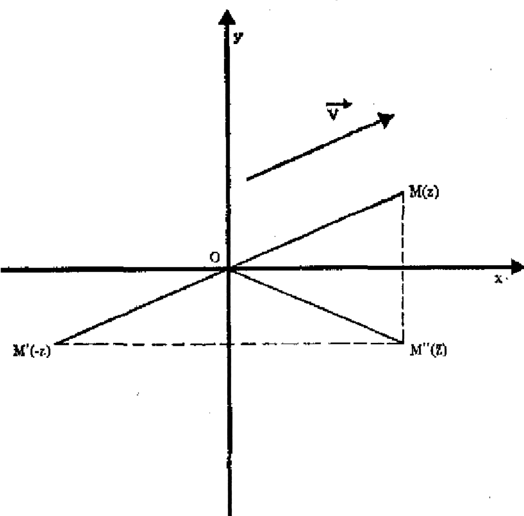


FIG. 11

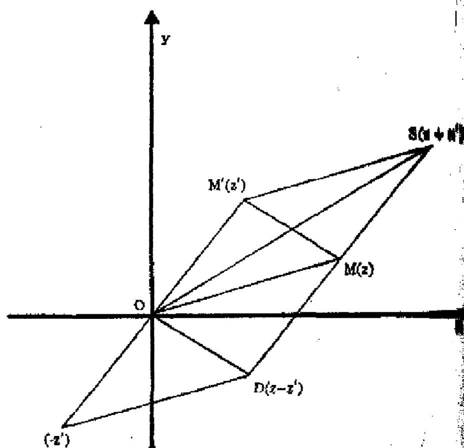


FIG. 12

En particular, a z y $-z$ corresponden $\vec{V}$ y $-\vec{V}$, o dos puntos M y M' simétricos respecto a O. Si los ejes son rectangulares a z y $\bar{z}$[22] corresponden dos puntos simétricos respecto al eje real (fig. 11):

Si M, M', D tienen por afijos respectivos z, z', $z-z'$ se tiene

$$\overrightarrow{OD} = \overrightarrow{OM} - \overrightarrow{OM'} = \overrightarrow{M'M}$$

por consiguiente: *El afijo del vector* M'M *es el afijo del extremo disminuido del afijo del origen.* La restricción al eje real nos determina el resultado bien conocido

$$\overline{M'M} = \overline{OM} - \overline{OM'}.$$

Además[23], para todo número real k

$$(2) \qquad (z \to \vec{V}) \Rightarrow (kz \to k\vec{V}).$$

EJERCICIO

Si a, b, z son complejos, demostrar que $(z-a)/(z-b)$ es real, si y sólo si existen dos números reales α y β verificando $\alpha + \beta = 1$ tales que $z = \alpha a + \beta b$. Interpretación geométrica; ¿cuál es el afijo del punto medio del segmento AB?

118. Interpretación de los números complejos mediante vectores o puntos del plano euclídeo orientado. Argumento de un número complejo

a) **C** provisto de la distancia $d(z, z') = |z - z'|$ es un espacio métrico (§ 116, *a*), luego si escribimos

$$\overrightarrow{OM} = \vec{V} = x\overrightarrow{OA} + y\overrightarrow{OB}, \qquad \overrightarrow{OM'} = \vec{V'} = x'\overrightarrow{OA} + y'\overrightarrow{OB}$$

la aplicación de $\mathbf{R}^2 \times \mathbf{R}^2$ en $\mathbf{R}_+$ definida por

$$(M, M') \to d(M, M') = |z - z'| = \sqrt{(x-x')^2 + (y-y')^2}$$

es una distancia llamada *distancia euclídea* (§ 110, ej. 1); $\mathbf{R}^2$ provisto de esta distancia se llama *plano euclídeo.* Pondremos

$$\|\vec{V}\| = \|\overrightarrow{OM}\| = d(0, M) = |z| = \sqrt{x^2 + y^2}$$

$\|\vec{V}\|$ se llama *norma euclídea* de V (ver § 231, *c*), las propiedades del módulo de z demuestran que (α real)

$$\|\vec{V}\| = 0 \Rightarrow \vec{V} = 0$$

$$\|\alpha\vec{V}\| = |\alpha|\,\|\vec{V}\|$$

$$\|\vec{V} + \vec{V'}\| \leq \|\vec{V}\| + \|\vec{V'}\|.$$

---

(22) Hemos dibujado las figuras 11 y 12 en ejes rectangulares, pero en este párrafo, salvo para esta interpretación de las imágenes z y $\bar{z}$, esto no es de ningún modo necesario.

(23) Las propiedades (1) y (2) nos permitirán decir en el capítulo 7 que el espacio vectorial $\mathbf{R} \times \mathbf{R}$ (sobre $\mathbf{R}$) es isomorfo a $\mathbf{C}$ espacio vectorial sobre $\mathbf{R}$ (§ 125, ej. 4).

El espacio de estos vectores libres de $\mathbf{R}^2$ provisto de esta norma es el espacio euclídeo de dos dimensiones (ver capítulo 15, § 231).

Definiremos en este espacio el *producto escalar* $\vec{V} \cdot \vec{V}'$ y diremos que $\vec{V}$ y $\vec{V}'$ no nulos son ortogonales si y sólo si $\vec{V} \cdot \vec{V}' = 0$. Si $\overrightarrow{OA}$ y $\overrightarrow{OB}$ son ortogonales y si $\|\overrightarrow{OA}\| = \|\overrightarrow{OB}\| = 1$

$$\vec{V} \cdot \vec{V}' = xx' + yy'$$

se dice entonces que los vectores libres $\overrightarrow{OA}$, $\overrightarrow{OB}$ forman una *base ortonormal* del espacio euclídeo de dos dimensiones. Vemos que $\|\vec{V}\| = \sqrt{\vec{V} \cdot \vec{V}}$ no es entonces más que la "*longitud*" de $\vec{V}$ de los estudios elementales y $d(\mathrm{M}, \mathrm{M}')$ la longitud del segmento MM'; la tercera propiedad de la distancia es la propiedad clásica de los triángulos, de donde su nombre de *desigualdad triangular*.

Definiremos en el § 235 los *ángulos* de dos vectores $(\vec{V}, \vec{V}')$ o de dos semirrectas (d, d'). La *medida* de un tal ángulo es un número real módulo 2π (§ 18, ej. 4 y la tabla), es decir, un elemento del grupo cociente $\mathbf{R}/2\pi\mathbf{Z}$ (§ 75, ej. 2); esta medida no es, pues, un número real, sino una clase módulo 2π y debería estar representada por $\dot{\alpha}$; por abuso de lenguaje diremos, sin embargo, que $\alpha, \alpha', \ldots$, representantes de $\dot{\alpha}$ son *medidas* del ángulo considerado; en fin, para simplificar la escritura designaremos por la misma notación[22] el ángulo y sus medidas; escribiremos, pues,

$$(\vec{V}, \vec{V}') \equiv \alpha \equiv \alpha' \qquad (\mathrm{mod}\ 2\pi).$$

Cuando para todo vector libre de $\mathbf{R}^2$: $\vec{V} = \overrightarrow{OM} = x\overrightarrow{OA} + y\overrightarrow{OB}$, tenemos

$$\|\overrightarrow{OA}\| = \|\overrightarrow{OB}\| = 1, \qquad \overrightarrow{OA} \cdot \overrightarrow{OB} = 0, \qquad (\overrightarrow{OA}, \overrightarrow{OB}) \equiv \frac{\pi}{2} \quad (\mathrm{mod}\ 2\pi)$$

diremos que $\overrightarrow{OA}$, $\overrightarrow{OB}$ es una *base ortonormal de sentido directo* (ver § 168), lo que supondremos en todo lo que sigue.

b) Sea $z = x + iy$ de imagen M, el *módulo* de z es

$$|z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2} = \|\overrightarrow{OM}\| = r$$

(22) De hecho harían falta cuatro notaciones (ver § 235):

— una para el par de vectores $(\vec{V}, \vec{V}')$,

— una para el ángulo de dos vectores $\widehat{(\vec{V}, \vec{V}')}$,

— una para la medida del ángulo α (elemento de $\mathbf{R}/2\pi\mathbf{Z}$),

— otra para los representantes de esta clase (elementos de $\mathbf{R}$).

es la *"longitud"* del vector libre $\vec{V} = \overrightarrow{OM}$; el *argumento* de $z \neq 0$ es la *medida* del ángulo $(\overrightarrow{OA}, \overrightarrow{OM})$

$$(\overrightarrow{OA}, \overrightarrow{OM}) \equiv \theta \qquad (2\pi).$$

Así, para dos números complejos no nulos de módulos y argumentos respectivos r y θ y r' y θ'

$$z = z' \Rightarrow [(r = \dot{r}, \ \text{y} \quad \theta \equiv \theta' \quad (\text{mod } 2\pi)].$$

Sea z un número complejo *no nulo,* existe un número complejo de módulo 1 único u tal que $z = |z| u$ (§ 116, *b*). Los números complejos de módulo 1 tienen por imagen los puntos del círculo de centro O y de radio 1 (*"círculo trigonométrico"*); luego si M es la imagen de $z \neq 0$, la semirrecta OM corta este círculo en un punto único U de afijo u y

$$\overrightarrow{OM} = \|\overrightarrow{OM}\| \overrightarrow{OU} \qquad z = |z| u$$

las coordenadas de U son $\cos \theta$ y $\operatorname{sen} \theta$, de donde

$$u = \cos \theta + i \operatorname{sen} \theta, \qquad z = r (\cos \theta + i \operatorname{sen} \theta)$$

con $(x^2 + y^2 \neq 0)$.

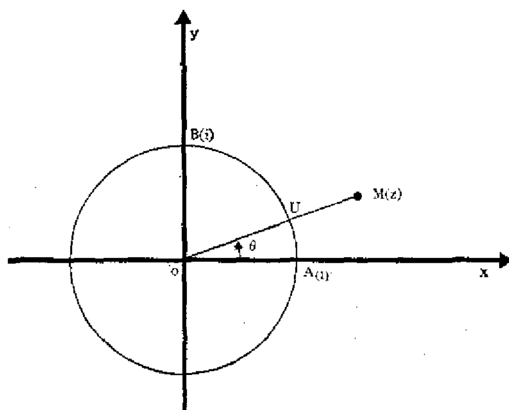


FIG. 13

$$\left\{\begin{array}{ll} x = r \cos \theta & (1) \\ y = r \operatorname{sen} \theta & (1') \end{array}\right. \Leftrightarrow \left\{\begin{array}{ll} r = \sqrt{x^2 + y^2} & (2) \\ \cos \theta = \dfrac{x}{\sqrt{x^2 + y^2}}, \quad \operatorname{sen} \theta = \dfrac{y}{\sqrt{x^2 + y^2}} & (2') \end{array}\right.$$

$z = x + iy$ es la *forma algebraica* de z, $z = r$ (cos $\theta + i$ sen θ) su *forma trigonométrica,* las fórmulas precedentes permiten pasar de una a otra.

Observemos que las fórmulas (2′) determinan el real θ módulo 2π, luego ($k \in \mathbf{Z}$)

$$r\,(\cos\theta + i\,\text{sen}\,\theta) = r'\,(\cos\theta' + i\,\text{sen}\,\theta') \Leftrightarrow \begin{cases} r' = r \\ \theta' = \theta + 2k\pi. \end{cases}$$

Consideremos finalmente la fórmula

$$\text{tg}\,\theta = \frac{y}{x}$$

válida para $x \neq 0$, a menudo útil, pero que no determina el argumento de $z = x + iy$, θ está definido por esta fórmula módulo π.

Observemos, en fin, que el *argumento de $z = 0$ no está definido.*

EJEMPLOS

z no es nulo:

1. Si z es el número real x

$$x < 0 \quad |x| = -x \quad \arg x \equiv \pi \quad (\text{mod } 2\pi).$$
$$x > 0 \quad |x| = x \quad \arg x \equiv 0 \quad (\text{mod } 2\pi).$$

2. Si z es el complejo puro iy, el módulo es $|y|$ y el argumento es congruente módulo 2π a $\frac{\pi}{2}$ o a $\frac{3\pi}{2}$ según que $y > 0$ o $y < 0$.

Así, $|i| = 1$, $\arg i \equiv \frac{\pi}{2}$ (mod 2π).

3. $|-z| = |z| \quad \arg(-z) \equiv \arg z + \pi \quad (\text{mod } 2\pi)$.

4. $|\bar{z}| = |z| \quad \arg(\bar{z}) \equiv -\arg z \quad (\text{mod } 2\pi)$.

5. Supongamos, en fin, siendo a y α dos números reales ($a \neq 0$),

$$z = a\,(\cos\alpha + i\,\text{sen}\,\alpha)$$
$$a > 0 \quad |z| = a \quad \arg z \equiv \alpha \quad (\text{mod } 2\pi)$$
$$a < 0 \quad |z| = -a \quad \arg z \equiv \alpha + \pi \quad (\text{mod } 2\pi).$$

c) Sea dos números complejos

$$z = r\,(\cos\theta + i\,\text{sen}\,\theta), \qquad z' = r'\,(\cos\theta' + i\,\text{sen}\,\theta').$$

Tendremos

$$zz' = rr'\,(\cos\theta\cos\theta' - \text{sen}\,\theta\,\text{sen}\,\theta') + i\,(\cos\theta\,\text{sen}\,\theta' + \text{sen}\,\theta\cos\theta').$$

$$\boxed{zz' = rr'\,[\cos(\theta + \theta') + i\,\text{sen}\,(\theta + \theta')]}$$

De donde por inducción

$$\prod_{k=1}^{n} r_k (\cos \theta_k + i \operatorname{sen} \theta_k) = \left[\prod_{k=1}^{n} r_k\right]\left[\cos \sum_{k=1}^{n} \theta_k + i \operatorname{sen} \sum_{k=1}^{n} \theta_k\right].$$

El módulo del producto de un número finito de números complejos es igual al producto de sus módulos, su argumento es congruente módulo 2π *a la suma de sus argumentos.*

En particular (n entero > 0),

$$[r (\cos \theta + i \operatorname{sen} \theta)]^n = r^n (\cos n\theta + i \operatorname{sen} n\theta).$$

Si $z \neq 0$

$$z^{-1} = \frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{r (\cos \theta - i \operatorname{sen} \theta)}{r^2}.$$

$$\boxed{\frac{1}{z} = \frac{1}{r} (\cos \theta - i \operatorname{sen} \theta)}$$

y siempre con $z \neq 0$

$$\frac{z'}{z} = \frac{r'}{r} [\cos (\theta' - \theta) + i \operatorname{sen} (\theta' - \theta)]$$

en particular, si $n = -n'$ es un entero negativo

$$z^n = z^{-n'} = \frac{1}{z^{n'}} = \frac{1}{r^{n'} (\cos n'\theta + i \operatorname{sen} n'\theta)}$$
$$= r^{-n'} (\cos n'\theta - i \operatorname{sen} n'\theta) = r^n (\cos n\theta + i \operatorname{sen} n\theta).$$

Tenemos, pues, para n entero racional cualquiera (fórmula de MOIVRE)

$$\boxed{(\cos \theta + i \operatorname{sen} \theta)^n = \cos n\theta + i \operatorname{sen} n\theta, \ n \in \mathbf{Z}}$$

d) Se verá en el tomo II (Análisis) que

$$e^{i\theta} = \cos \theta + i \operatorname{sen} \theta$$

en particular (k entero racional),

$$e^{2ik\pi} = 1 \qquad e^{i\pi} = -1$$

tendremos ($n \in \mathbf{Z}$)

$$e^{i\theta} e^{i\theta'} = e^{i(\theta + \theta')}, \qquad (e^{i\theta})^{-1} = e^{-i\theta}, \qquad (e^{i\theta})^n = e^{ni\theta}.$$

Todo número complejo $z = r$ (cos $\theta + i$ sen θ) puede entonces escribirse bajo la forma

$$z = re^{i\theta}$$

llamada *forma exponencial,* particularmente cómoda gracias a las fórmulas recordadas anteriormente para efectuar productos y cocientes.

EJERCICIOS

1. Encontrar el módulo y el argumento de los números complejos siguientes (α y β reales)

$$1 + i, \quad 1 - i\sqrt{3}, \quad 3 + i\sqrt{3}, \quad 2\alpha + i(1 - \alpha^2)$$

$$1 + \varepsilon \cos \alpha + \varepsilon' i \operatorname{sen} \alpha, \quad 1 + \varepsilon \operatorname{sen} \alpha + \varepsilon' i \cos \alpha \quad (\varepsilon^2 = \varepsilon'^2 = 1)$$

$$\frac{1 + \cos \alpha + i \operatorname{sen} \alpha}{1 - \cos \alpha - i \operatorname{sen} \alpha}, \quad (1 + i\sqrt{3})^6, \quad (1 - i)^n \quad (n \in \mathbf{Z})$$

$$1 + i \operatorname{tg} \alpha, \quad \operatorname{tg} \alpha + i \operatorname{tg} \beta, \quad \frac{e^{i\alpha} + e^{i\beta}}{1 + e^{i(\alpha+\beta)}}.$$

2. Si p y q son dos enteros naturales, encontrar z para que z^p y z^q sean conjugados.

119. Raíces n-ésimas de un número complejo

a) Sea z el número complejo r (cos $\theta + i$ sen θ) $\neq 0$, busquemos

$$z' = r' (\cos \theta' + i \operatorname{sen} \theta')$$

tal que para n entero > 0 sea

$$z'^n = z$$

$[r' (\cos \theta' + i \operatorname{sen} \theta')]^n = r'^n (\cos n\theta' + i \operatorname{sen} n\theta') = r (\cos \theta + i \operatorname{sen} \theta)$ da (con k entero racional), puesto que r y r' son positivos,

$$\begin{cases} r^{n\prime} = r \\ n\theta' = \theta + 2k\pi \end{cases} \Leftrightarrow \begin{cases} r' = \sqrt[n]{r} \\ \theta' = \dfrac{\theta}{n} + \dfrac{2k\pi}{n}. \end{cases}$$

Todos los números obtenidos tienen el mismo módulo, serán distintos si sus argumentos no son congruentes módulo 2π: luego para obtenerlos todos es necesario y suficiente dar a k, n valores enteros consecutivos; por ejemplo, 0, 1, 2, ..., $n-1$:

TEOREMA. — *Todo número complejo* $z = r$ (cos $\theta + i$ sen θ) *no nulo tiene* n *raíces* n-ésimas

$$\boxed{\begin{gathered} z_k = \sqrt[n]{r}\left(\cos \frac{\theta + 2k\pi}{n} + i \operatorname{sen} \frac{\theta + 2k\pi}{n} \right) \\ 0 \leq k \leq n - 1 \end{gathered}}.$$

Se tiene, pues,

$$|z_{k+1}| = |z_k| \qquad \arg z_{k+1} \equiv \arg z_k + \frac{2\pi}{n} \qquad (\text{mod } 2\pi)$$

se pasa, pues, de la imagen M_k a la de M_{k+1} por una *rotación* de centro O y de ángulo $\frac{2\pi}{n}$, luego:

*Las imágenes de las n raíces n-*ésimas *de un número complejo no nulo para $n > 2$, son los vértices de un polinomio regular de n lados con centro en* O.

b) Volvamos sobre el caso particular de las raíces cuadradas utilizando la forma trigonométrica $z = r (\cos \theta + i \text{ sen } \theta) \neq 0$.

Estas dos raíces son

$$z_0 = r \left(\cos \frac{\theta}{2} + i \text{ sen } \frac{\theta}{2} \right),$$

$$z_1 = r \left[\cos \left(\frac{\theta}{2} + \pi \right) + i \text{ sen } \left(\frac{\theta}{2} + \pi \right) \right] = -z_0$$

encontramos que son opuestas y que, por consiguiente, sus imágenes son simétricas respecto a O.

120. Raíces n-ésimas de la unidad

a) Si Z' es una raíz n-ésima de $z = r (\cos \theta + i \text{ sen } \theta) \neq 0$ todas las raíces n-ésimas z' de z son tales que

$$z'^n = Z'^n = z$$

es decir,

$$\left(\frac{Z'}{z'} \right)^n = 1, \qquad z' = Z'\omega_k$$

donde ω_k es una de las n raíces n-ésimas de la unidad, por consiguiente:

TEOREMA. — *Se obtienen las n raíces n-*ésimas *de un número complejo no nulo multiplicando una de ellas por las n raíces n-*ésimas *de la unidad.*

b) Es suficiente, en consecuencia, estudiar las n raíces n-ésimas de la unidad; tendremos, al ser el módulo y el argumento de 1, respectivamente, 1 y 0 (módulo 2π)

$$\boxed{\omega_k = \cos \frac{2k\pi}{n} + i \text{ sen } \frac{2k\pi}{n} = e^{\frac{2ik\pi}{n}} \qquad (0 \leq k \leq n-1)}\,.$$

Observemos las fórmulas $\overline{\omega_k} = (\omega_k)^{-1} = \omega_{-k}$ (la primera igualdad traduce la propiedad característica de los complejos de módulo 1: $\bar{u} = u^{-1}$ (§ 116, *b*)).

EJEMPLOS

1. $n = 2$ $\quad \omega^2 = 1$ $\quad$ da $\quad 1$ y -1.
2. $n = 3$ $\quad \omega^3 = 1$ $\quad$ da

$$\omega_0 = 1 \qquad \omega_1 = \cos\frac{2\pi}{3} + i \operatorname{sen}\frac{2\pi}{3} = \frac{-1 + i\sqrt{3}}{2} = j^{(23)}$$

$$\omega_2 = \cos\frac{4\pi}{3} + i \operatorname{sen}\frac{4\pi}{3} = \frac{-1 - i\sqrt{3}}{2} = j^2 = \bar{j}.$$

Observemos en particular que (fórmula que es general, ver ej. 3)

$$1 + j + j^2 = 0.$$

3. $n = 4$ $\quad \omega^4 = 1$ $\quad$ da $\quad \omega_0 = 1 \quad \omega_1 = i \quad \omega_2 = -1 \quad \omega_3 = -i.$

4. $n = 2p$: $\omega_k = \cos\dfrac{2k\pi}{2p} + i \operatorname{sen}\dfrac{2k\pi}{2p} = \cos\dfrac{k\pi}{p} + i \operatorname{sen}\dfrac{k\pi}{p}$

hay dos y sólo dos números reales $\omega_0 = 1$, $\omega_p = -1$; respecto a los $n - 2 = 2(p - 1)$ restantes, podemos agruparles por pares de raíces conjugadas ω_k y $\omega_{-k} = \overline{\omega_k}$, con $1 \leq k \leq p - 1$, las imágenes son los vértices de un polígono regular teniendo dos vértices sobre el eje real.

5. $n = 2p + 1$: $\omega_k = \cos\dfrac{2k\pi}{2p + 1} + i \operatorname{sen}\dfrac{2k\pi}{2p + 1}$.

una sola es real $\omega_0 = 1$, las $n - 1 = 2p$ restantes se agrupan por pares de raíces conjugadas ω_k y $\omega_{-k} = \overline{\omega_k}$ con $1 \leq k \leq p$, las imágenes son los vértices de un polígono regular siempre simétrico respecto al eje real, pero teniendo sólo un vértice encima.

c) Las n raíces n-ésimas de la unidad describen un subgrupo del grupo multiplicativo **U**; en efecto, cualquiera que sean k y k' enteros

$$e^{\frac{2ik\pi}{n}} \left(e^{\frac{2ik'\pi}{n}} \right)^{-1} = e^{\frac{2i(k-k')\pi}{n}}$$

luego $\omega_k(\omega_{k'})^{-1}$ es una raíz n-ésima de 1, su conjunto es un subgrupo de **U** (§ 72, *c*). Designémosle por $\mathbf{U}_n$ y consideremos la aplicación de $\mathbf{U}_n$ en $\mathbf{Z}/n\mathbf{Z}$ definida por

$$e^{\frac{2ik\pi}{n}} \to k$$

esta aplicación es biyectiva, pues

$$k = k' \Leftrightarrow k = k' \pmod{n} \Leftrightarrow e^{\frac{2ik\pi}{n}} = e^{\frac{2ik'\pi}{n}}.$$

(23) Los físicos (en particular los electricistas), que utilizan los números complejos, reservan i para la intensidad de corriente, mientras que designan por j una de las raíces cuadradas de -1; utilizan otra letra, en general ω, para representar las raíces cúbicas de 1. En los libros de electricidad encontramos $j^2 = -1$ y $\omega^3 = 1$.

Por otra parte, cualquiera que sean k y k'

$$e^{\frac{2ik\pi}{n}} e^{\frac{2ik'\pi}{n}} = e^{\frac{2i(k+k')\pi}{n}} \Rightarrow \left(\dot{\widehat{k+k'}}\right) = \dot{k} + \dot{k}'$$

de donde (ver § 75):

Teorema. — *El conjunto de las n raíces n-ésimas de la unidad es un subgrupo del grupo multiplicativo* $\mathbf{U}$ *de los números complejos de módulo* 1; *es isomorfo al grupo aditivo* $\mathbf{Z}/n\mathbf{Z}$.

La segunda parte de este teorema ha sido ya demostrada de una manera general en el § 83 (pues $\mathbf{U}_n$ es un grupo monógeno de orden n). Busquemos en qué condiciones $e^{\frac{2ip\pi}{n}}$ engendra $\mathbf{U}_n$. Es necesario y suficiente que cualquiera que sea el entero q, exista un entero u tal que

$$e^{\frac{2iq\pi}{n}} = \left(e^{\frac{2ip\pi}{n}}\right)^u = e^{\frac{2ipu\pi}{n}}$$



es decir, que existe también un entero v tal que

$$\frac{2q}{n} = \frac{2pu}{n} + 2v \Leftrightarrow q = pu + nv$$

$pu + nv$ describe el ideal (p, n) de $\mathbf{Z}$ (§ 99); es necesario, pues, que este ideal sea $\mathbf{Z}$, es decir, que p y n tengan por m. c. d. 1, esto es, que sean primos entre sí (§ 99, teorema 5).

Teorema y definición. — *Un elemento* $e^{\frac{2ip\pi}{n}}$ *del grupo* $\mathbf{U}_n$ *de las raíces n-ésimas de la unidad, engendra este grupo si y sólo si n y p son primos entre sí. Se dice entonces que* $e^{\frac{2ip\pi}{n}}$ *es una* raíz primitiva n-ésima *de* 1.

Así, $\omega_1 = e^{\frac{2i\pi}{n}}$ es siempre raíz primitiva n-ésima de 1 y para n primo toda raíz n-ésima distinta de $\omega_0 = 1$ es primitiva.

Para $n = 6$, ω_1, ω_5 son primitivas, pero no ω_0, ω_2, ω_3, ω_4.

EJERCICIOS

1. Encontrar las raíces n-ésimas de z para los valores siguientes de n y de z

$$z = 4ab + 2(a^2 - b^2)i \quad (a,\ b \text{ reales}) \quad n = 2$$

$$z = 1 + i, \quad z = 1 - i \quad n = 2$$

$$z = 1 + i, \quad z = \frac{1 + i\sqrt{3}}{1 - i\sqrt{3}} \quad n = 3$$

$$z = \frac{1 + ia}{1 - ia} \quad (a \text{ real}) \quad n \in \mathbf{N}$$

(se pondrá $a = \operatorname{tg} \alpha$ y se demostrará que las raíces n-ésimas son de la forma $\dfrac{1 + i \operatorname{tg} x}{1 - i \operatorname{tg} x}$, x real).

2. Encontrar las órdenes de cada uno de los elementos de U_{30} (ver § 83, ej. 4).
3. Calcular

$$A_{np} = \sum_{k=0}^{n-1} \left(e^{\frac{2ip\pi}{n}} \right)^k$$

observar en particular la fórmula

$$\sum_{k=0}^{n-1} e^{\frac{2ik\pi}{n}} = 0.$$

III. Aplicaciones de los números complejos

121. Aplicaciones a los cálculos trigonométricos

a) Si n es un entero natural, calculemos cos na y sen na en función de cos a y sen a; la fórmula de MOIVRE

$$\cos na + i \operatorname{sen} na = (\cos a + i \operatorname{sen} a)^n$$

nos da con la ayuda de la fórmula del binomio

$$(1) \quad \cos na = (\cos a)^n - C_n^2 (\cos a)^{n-2} (\operatorname{sen} a)^2 + \dots + (-1)^p C_n^{2p} (\cos a)^{n-2p} (\operatorname{sen} a)^{2p} + \dots$$

$$(2) \quad \operatorname{sen} na = C_n^1 (\cos a)^{n-1} \operatorname{sen} a - C_n^3 (\cos a)^{n-3} (\operatorname{sen} a)^3 + \dots + (-1)^p C_n^{2p+1} (\cos a)^{n-2p-1} (\operatorname{sen} a)^{2p+1} + \dots$$

naturalmente, se trata de sumas finitas; el último término no está indicado, porque su forma depende de la paridad de n. Estas fórmulas se las puede llamar "*fórmulas homogéneas*", pues dan cos na y sen na en la forma de polinomios homogéneos (ver § 186, *b*) en cos a y sen a.

Si cos na y cos a son no nulos, se obtiene

$$(3) \quad \operatorname{tg} na = \frac{C_n^1 \operatorname{tg} a - C_n^3 (\operatorname{tg} a)^3 + \dots + (-1)^p C_n^{2p+1} (\operatorname{tg} a)^{2p+1} + \dots}{1 - C_n^2 (\operatorname{tg} a)^2 + \dots + (-1)^p C_n^{2p} (\operatorname{tg} a)^{2p} + \dots}.$$

Así[24],

$(1')$ $\cos 2a = \cos^2 a - \operatorname{sen}^2 a$ $\qquad$ $(1'')$ $\cos 3a = \cos^3 a - 3 \cos a \operatorname{sen}^2 a$

$(2')$ $\operatorname{sen} 2a = 2 \operatorname{sen} a \cos a$ $\qquad$ $(2'')$ $\operatorname{sen} 3a = 3 \operatorname{sen} a \cos^2 a - \operatorname{sen}^3 a$

$(3')$ $\operatorname{tg} 2a = \dfrac{2 \operatorname{tg} a}{1 - \operatorname{tg}^2 a}$ $\qquad$ $(3'')$ $\operatorname{tg} 3a = \dfrac{3 \operatorname{tg} a - \operatorname{tg}^3 a}{1 - 3 \operatorname{tg}^2 a}.$

(24) Utilizamos el abuso de notación clásica: $\cos^2 a$ por $(\cos a)^2$, $\operatorname{sen}^3 a$ por $(\operatorname{sen} a)^3$, etc.

Es interesante buscar si, gracias a la fórmula $\cos^2 a + \operatorname{sen}^2 a = 1$, se puede obtener cos na y sen na en la forma de un polinomio en cos a o en sen a, se verá fácilmente que cualquiera que sea n

$$\cos na = \mathrm{P}\ (\cos a)$$

y que

$$\begin{array}{ll} n \text{ par} & \cos na = \mathrm{Q}\ (\operatorname{sen} a) \\ n \text{ impar} & \operatorname{sen} na = \mathrm{R}\ (\operatorname{sen} a) \end{array}$$

P, Q, R son polinomios de grado n (ver ej. 1), así

$$\cos 2a = 2 \cos^2 a - 1 = 1 - 2 \operatorname{sen}^2 a$$
$$\cos 3a = 4 \cos^3 a - 3 \cos a \qquad \operatorname{sen} 3a = 3 \operatorname{sen} a - 4 \operatorname{sen}^3 a.$$

b) Es a menudo cómodo transformar un polinomio f en cos a y sen a en *polinomio trigonométrico*, es decir, determinar las sucesiones finitas (a_h) y (b_h) tales que

$$f\ (\cos a, \operatorname{sen} a) = \sum_{h=0}^{n} (a_h \cos ha + b_h \operatorname{sen} ha)$$

se puede hacer por aplicación repetida de las fórmulas elementales de transformaciones de "productos en sumas". Se simplifica los cálculos utilizando al mismo tiempo las fórmulas

$$\cos^2 a = \frac{1 + \cos 2a}{2} \qquad \operatorname{sen}^2 a = \frac{1 - \cos 2a}{2}.$$

Los números complejos nos permiten obtener sistemáticamente esta descomposición. Observemos primero que las fórmulas

$$e^{ia} = \cos a + i \operatorname{sen} a \qquad e^{-ia} = \cos a - i \operatorname{sen} a$$

nos dan las fórmulas de EULER

$$\cos a = \frac{e^{ia} + e^{-ia}}{2} \qquad \operatorname{sen} a = \frac{e^{ia} - e^{-ia}}{2i};$$

y se deduce

$$(\cos a)^n = \left(\frac{e^{ia} + e^{-ia}}{2} \right)^n \qquad (\operatorname{sen} a)^n = \left(\frac{e^{ia} - e^{-ia}}{2i} \right)^n$$

que al desarrollar por la fórmula del binomio, agrupar los términos equidistantes de los extremos y aplicando las fórmulas de EULER, nos determina un polinomio trigonométrico (ver ej. 2); por ejemplo,

$$\begin{aligned} 2^3 \cos^3 a &= (e^{ia} + e^{-ia})^3 = e^{3ia} + 3e^{2ia}e^{-ia} + 3e^{ia}e^{-2ia} + e^{-3ia} \\ &= (e^{3ia} + e^{-3ia}) + 3(e^{ia} + e^{-ia}) = 2 \cos 3a + 6 \cos a \\ (2i)^3 \operatorname{sen}^3 a &= (e^{ia} - e^{-ia})^3 = e^{3ia} - 3e^{2ia}e^{-ia} + 3e^{ia}e^{-2ia} - e^{-3ia} \\ &= (e^{3ia} - e^{-3ia}) - 3(e^{ia} - e^{-ia}) = 2i \operatorname{sen} 3a - 6i \operatorname{sen} a \end{aligned}$$

de donde

$$\cos^3 a = \frac{1}{4}\cos 3a + \frac{3}{4}\cos a, \qquad \operatorname{sen}^3 a = -\frac{1}{4}\operatorname{sen} 3a + \frac{3}{4}\operatorname{sen} a.$$

Este método se aplica a todo polinomio en cos a y sen a (ver ej. 3).

c) Factoricemos las dos sumas C y S siguientes (α y β números reales, n entero natural)

$$C = \cos\alpha + \cos(\alpha+\beta) + \ldots + \cos[\alpha + (n-1)\beta]$$
$$S = \operatorname{sen}\alpha + \operatorname{sen}(\alpha+\beta) + \ldots + \operatorname{sen}[\alpha + (n-1)\beta]$$

para esto calculemos $C + iS$

$$C + iS = e^{i\alpha} + e^{i(\alpha+\beta)} + \ldots + e^{i[\alpha+(n-1)\beta]} = e^{i\alpha}\sum_{k=0}^{n-1} e^{ik\beta}$$

es una progresión geométrica de n términos, de primer término $a = e^{i\alpha}$ y de razón $q = e^{i\beta}$, luego si $q \neq 1$ (es decir, $\beta \neq 2h\pi$, h entero), tendremos poniendo $\beta = 2\beta'$

$$C + iS = a(1 + q + \ldots + q^{n-1}) = a\frac{1-q^n}{1-q} = e^{i\alpha}\frac{1-\cos n\beta - i\operatorname{sen} n\beta}{1-\cos\beta - i\operatorname{sen}\beta}$$
$$= e^{i\alpha}\frac{2\operatorname{sen}^2 n\beta' - 2i\operatorname{sen} n\beta'\cos n\beta'}{2\operatorname{sen}^2\beta' - 2i\operatorname{sen}\beta'\cos\beta'} = \frac{\operatorname{sen} n\beta'}{\operatorname{sen}\beta'}e^{i\alpha}\frac{\operatorname{sen} n\beta' - i\cos n\beta'}{\operatorname{sen}\beta' - i\cos\beta'}$$
$$= \frac{\operatorname{sen} n\beta'}{\operatorname{sen}\beta'}e^{i\alpha}\frac{\cos n\beta' + i\operatorname{sen} n\beta'}{\cos\beta' + i\operatorname{sen}\beta'} = \frac{\operatorname{sen} n\beta'}{\operatorname{sen}\beta'}e^{i[\alpha+(n-1)\beta']}$$

de donde

$$C = \frac{\operatorname{sen}\frac{n\beta}{2}}{\operatorname{sen}\frac{\beta}{2}}\cos\left[\alpha + (n-1)\frac{\beta}{2}\right], \qquad S = \frac{\operatorname{sen}\frac{n\beta}{2}}{\operatorname{sen}\frac{\beta}{2}}\operatorname{sen}\left[\alpha + (n-1)\frac{\beta}{2}\right]$$

en particular, para x real y n entero natural

$$1 + \cos x + \cos 2x + \ldots + \cos nx = \frac{\operatorname{sen}(n+1)\frac{x}{2}}{\operatorname{sen}\frac{x}{2}}\cos\frac{n}{2}x$$

$$\operatorname{sen} x + \operatorname{sen} 2x + \ldots + \operatorname{sen} nx = \frac{\operatorname{sen}(n+1)\frac{x}{2}}{\operatorname{sen}\frac{x}{2}}\operatorname{sen}\frac{n}{2}x.$$

EJERCICIOS

1. Demostrar que los polinomios P, Q, R definidos anteriormente son de grado n (calcular el término de mayor grado utilizando el ejercicio 4, § 92).
2. Efectuar los cálculos que dan cos na y sen na en la forma de polinomio trigonométrico (distinguir dos casos según la paridad de n).
3. Escribir en la forma de polinomio trigonométrico

$$2 \cos^4 x \operatorname{sen}^3 x + \cos^2 x \operatorname{sen}^2 x.$$

4. Calcular (con α, β ρ números reales, n entero natural)

$$\sum_{k=0}^{n-1} \rho^k \cos(\alpha + k\beta), \qquad \sum_{k=0}^{n-1} \rho^k \operatorname{sen}(\alpha + k\beta).$$

122. Aplicación de números complejos a la geometría afín plana

a) Consideremos la traslación $M \to M'$ definida por

$$\overrightarrow{MM'} = \vec{A} \tag{1}$$

a, z y z' son los afijos respectivos de $\vec{A}$, M, M′, tendremos

$$z' = z + a. \tag{1'}$$

Recíprocamente (1′) implica (1), luego a cada traslación corresponde biunívocamente un número complejo. Además, las notaciones evidentes

$$\begin{cases} \overrightarrow{MM'} = \vec{A} \\ \overrightarrow{M'M''} = \vec{B} \end{cases} \Rightarrow z'' = z + (a + b)$$

luego si t y s son traslaciones representadas, respectivamente, por los números complejos a y b, la traslación $s \circ t = t \circ s$ está representada por $a + b$, de donde: *El grupo de las traslaciones del plano es isomorfo al grupo aditivo de los números complejos.*

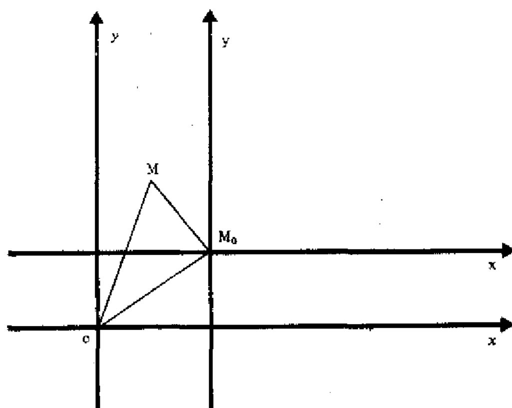


FIG. 14

b) Efectuemos sobre los ejes Ox y Oy la traslación representada por $z_0 = x_0 + iy_0$, que los traslada a M_0X, M_0Y, donde M_0 tiene por afijo z_0. Si $z = x + iy$ y $Z = X + iY$ son los afijos de M, respectivamente, respecto a Ox y Oy, y M_0X y M_0Y, tendremos

$$(2) \qquad \overrightarrow{OM} = \overrightarrow{OM_0} + \overrightarrow{M_0M}$$

$$(2') \qquad z = z_0 + Z \Leftrightarrow \begin{cases} x = x_0 + X \\ y = y_0 + Y. \end{cases}$$

En una traslación de ejes *"el afijo antiguo" de* M *es igual al "afijo nuevo" aumentado en el afijo antiguo del nuevo origen.*

c) Hemos visto en el § 117 que si k es real no nulo, con z y z' afijos respectivos de M y M'

$$(3) \qquad \overrightarrow{OM'} = k\overrightarrow{OM} \Leftrightarrow z' = kz \qquad (3')$$

la fórmula (3') representa, pues, la *homotecia* $H(0, k)$. Estudiemos la homotecia $H(M_0, k)$, y sea z_0 el afijo de M_0. Al considerar los ejes M_0X y M_0Y deducidos de Ox y Oy por la traslación z_0, tendremos

$$Z' = kZ \Leftrightarrow z' - z_0 = k(z - z_0).$$

Recíprocamente estudiemos la transformación $M(z) \rightarrow M'(z')$ definida por

$$(4) \qquad z' = az + b \qquad (a \in R^*, \ b \in C).$$

Si hay un punto invariante z_0 se tendrá $z_0 = az_0 + b$; luego existe un punto invariante y uno sólo si $a \neq 1$. En este caso (4) puede escribirse

$$z' - z_0 = a(z - z_0),$$

es la homotecia $H(M_0, k)$. Si $a = 1$, es la traslación asociada al número complejo b.

123. Aplicación de los números complejos a la geometría euclídea plana

a) Dado el número complejo $a = k\ (\cos\ \alpha + i\ \text{sen}\ \alpha) = ke^{i\alpha}$ (k real estrictamente positivo), interpretemos geométricamente la aplicación

$$(1) \qquad z \rightarrow z' = az.$$

Sea M y M' las imágenes respectivas de z y z', pongamos $z = re^{i\theta}$ $z' = r'e^{i\theta'}$, tendremos

$$r'e^{i\theta'} = ke^{i\alpha}re^{i\theta} \Leftrightarrow \begin{cases} r' = kr \\ \theta' = \theta + \alpha \qquad (\text{mod } 2\pi) \end{cases}$$

de donde

$$\begin{cases} \| \overrightarrow{OM'} \| = k \| \overrightarrow{OM} \| \\ (Ox, \overrightarrow{OM'}) \equiv (Ox, \overrightarrow{OM}) + \alpha \qquad (\text{mod } 2\pi) \end{cases}$$

luego (1) representa la semejanza S(0, k, α); se sabe que S(0, $-k$, $\alpha + \pi$) = S(0, k, α); luego si se impone a k el ser estrictamente positivo *a toda semejanza plana* S(0, k, α), *se le puede hacer corresponder de manera biunívoca el número complejo* $ke^{i\alpha}$. Además, si T(O, l, β) es otra semejanza de centro O, tendremos con las notaciones evidentes

$$\begin{cases} z' = ke^{i\alpha} z = az \\ z'' = le^{i\beta} z' = bz' \end{cases} \Rightarrow z'' = kle^{i(\alpha+\beta)} z = (ab)z.$$

En consecuencia, si dos semejanzas de centro O están representadas, respectivamente, por los números complejos a y b no nulos, la semejanza "producto" está representada por ab, es decir:

El conjunto de las semejanzas planas de centro O, *provisto de la composición de aplicaciones es un grupo isomorfo al grupo multiplicativo* **C***.

En particular, si $k = 1$, la fórmula

$$z' = e^{i\alpha}z$$

define la *rotación* R(0, α). Si $\alpha = 0$ o $\alpha = \pi$ y k real no nulo, volvemos a encontrar la homotecia H(0, k) definida por $z' = kz$.

Si $\mathcal{H}_+$, $\mathcal{H}$, $\mathcal{R}$, $\mathcal{S}$ representan, respectivamente, los conjuntos de las homotecias positivas, de las homotecias, de las rotaciones y de las similitudes todas de centro O, provisto de la composición de las aplicaciones, $\mathcal{H}_+$, $\mathcal{H}$, $\mathcal{R}$, $\mathcal{S}$ son los grupos isomorfos, respectivamente, de los grupos multiplicativos $\mathbf{R}_+^*$, $\mathbf{R}^*$, U, $\mathbf{C}^*$ y se tiene

$$\mathcal{H}_+ \subset \mathcal{H} \subset \mathcal{S}, \qquad \mathcal{R} \subset \mathcal{S}.$$

b) Considerando los ejes M_0X, M_0Y deducidos de Ox y Oy por la traslación z_0, tendremos para la semejanza S(M_0, k, α)

(2) $$Z' = ke^{i\alpha}Z \Leftrightarrow z' - z_0 = ke^{i\alpha}(z - z_0)$$

la fórmula (2) es de la forma

(3) $$z' = az + b.$$

Consideremos *a priori* la fórmula (3) en que a y b son números complejos cualesquiera. Si (3) se escribe en la forma (2), encontraremos z_0 tal que

(3') $$z_0 = az_0 + b$$

si y sólo si $a \neq 1$, en este caso la transformación definida por (3) admite un punto invariante único M_0 de afijo z_0, restando (3') de (3), se obtiene

$$z' - z_0 = a(z - z_0).$$

Si $a \neq 0$ y $a \neq 1$ poniendo $a = ke^{i\alpha}$ $(k > 0)$ se encuentra la forma (2). Si $a = 1$ volvemos a encontrar una traslación, de donde:

TEOREMA. — *La transformación del plano,* $M(z) \to M'(z')$ definida por

$$z' = az + b \qquad (a \neq 0)$$

es biyectiva.

Si $a = ke^{i\alpha} \neq 1$, $(k > 0)$ *es una semejanza de razón* k *y de ángulo* α; *se reduce a una rotación para* $k = 1$ *y a una homotecia para* $\alpha \equiv 0 \pmod{\pi}$.

Si $a = 1$, *es una traslación.*

OBSERVACION

El conjunto de las semejanzas del plano provisto de la ley de composición de las aplicaciones no describe un grupo; en efecto,

$$\begin{cases} z' = az + b \\ z'' = a'z' + b' \end{cases} \Rightarrow z'' = Az + B \qquad A = aa'.$$

Se puede tener $a \neq 1$, $a' \neq 1$, $aa' = 1$; por el contrario, el conjunto de las biyecciones definidas por $z' = az + b$ $(a \neq 0)$ es un grupo.

c) Demos a los ejes Ox, Oy una rotación R(O, α), se convierten en OX, OY; sea z y Z los afijos respectivos de un *mismo punto* M, respectivamente, respecto a xOy y XOY, tenemos

$$\begin{cases} |z| = |Z| = \|\overrightarrow{OM}\| \\ \arg z = (Ox, \overrightarrow{OM}) = (Ox, OX) + (OX, \overrightarrow{OM}) \equiv \alpha + \arg Z \qquad \pmod{2\pi} \end{cases}$$

de donde

$$z = Ze^{i\alpha}$$

es decir,

$$x + iy = (X + iY)(\cos\alpha + i \operatorname{sen}\alpha) \qquad \begin{cases} x = X\cos\alpha - Y\operatorname{sen}\alpha \\ y = X\operatorname{sen}\alpha + Y\cos\alpha. \end{cases}$$

d) Consideremos dos puntos A(a) y B(b) y un punto M(z) distinto de B

$$\left|\frac{z-a}{z-b}\right| = \frac{|z-a|}{|z-b|} = \frac{\|\overrightarrow{MA}\|}{\|\overrightarrow{MB}\|}$$

$$\arg\frac{z-a}{z-b} \equiv \arg(z-a) - \arg(z-b) \equiv (Ox, \overrightarrow{AM}) - (Ox, \overrightarrow{BM})$$

$$\equiv (\overrightarrow{BM}, \overrightarrow{AM}) = (\overrightarrow{MA}, \overrightarrow{MB}). \qquad \pmod{2\pi}$$

TEOREMA. — *El conjunto de los puntos* M(z) *tales que* $\left|\frac{z-a}{z-b}\right| = k > 0$ *es un círculo si* $k \neq 1$ *y una recta si* $k = 1$ *(la mediatriz de* A(*a*) B(*b*)).

El conjunto de los puntos M(z) *tales que* $\arg \frac{z-a}{z-b} \equiv \alpha$ (mod 2π) (resp. mod π) *es un arco de círculo* (resp. *un círculo*) *si* $\alpha \neq 0$ (mod π).

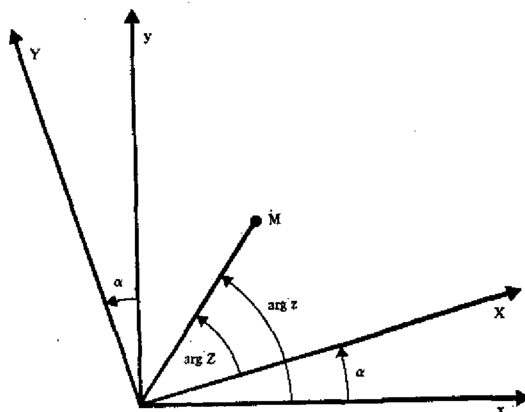


FIG. 15

EJERCICIOS

1. ¿Cuál es el conjunto de los puntos M(z) tales que $(z-a)/(z-b)$ sea real positivo (resp. real negativo)?

2. ¿Qué resultado da el estudio anterior (subpárrafo *d*) para $b = \bar{a}$? Relación con el ejercicio 2 del § 116.

124. Geometría del plano analítico $\tilde{C} = C \cup \{\infty\}$. Grupo circular

a) La aplicación $z \rightarrow z^{-1}$ no está definida más que en C^*: es una biyección de C^* sobre sí mismo; podemos "añadir" a C un "*punto del infinito*", representado ∞ y proveer al conjunto $\tilde{C} = C \cup \{\infty\}$ de una suma y una multiplicación no definidas totalmente, induciendo sobre C la adición y multiplicación de números complejos escribiendo

$$(\forall z \in C) \qquad z + \infty = \infty, \qquad (\forall z \in \tilde{C} - \{0\}) \qquad z\,\infty = \infty$$

y

$$\frac{1}{0} = \infty, \qquad \frac{1}{\infty} = 0.$$

Los elementos de C se llamarán *finitos*. Observemos que añadiendo ∞ a C *destruimos* la estructura de cuerpo e incluso la de grupo aditivo, ya que el elemento ∞ no es regular para la adición.

Sin embargo, vamos a demostrar que se puede definir una *biyección* de $\tilde{C}$ sobre una esfera **S** e incluso suministrar a $\tilde{C}$ de una distancia δ tal que la biyección anterior sea un *homeomorfismo*(25) de $\tilde{C}$, provisto de la distancia δ, sobre **S**, supuesta **S** sumergida en el espacio euclídeo $\mathbf{R}^3$ que posee la distancia euclídea habitual.

Examinemos las fórmulas del ejercicio del § 115 ($z = x + iy$)

$$(1) \qquad X = \frac{z+\bar{z}}{1+z\bar{z}} = \frac{2x}{1+x^2+y^2} \qquad Y = i\frac{\bar{z}-z}{1+z\bar{z}} = \frac{2y}{1+x^2+y^2}$$

$$Z = \frac{1-z\bar{z}}{1+z\bar{z}} = \frac{1-x^2-y^2}{1+x^2+y^2}$$

tenemos $X^2 + Y^2 + Z^2 = 1$ y observando que $Z \neq -1$

$$(1') \qquad \frac{X}{x} = \frac{Y}{y} = \frac{Z+1}{1} = \frac{2}{1+x^2+y^2}.$$

A todo punto m de afijo z de $\tilde{C}$ hacemos también corresponder un punto único $M(X, Y, Z)$ de la esfera **S** de centro O y de radio 1. Sea P el punto $(0, 0, -1)$ de **S**, las fórmulas (1′) muestran que P, m, M están alineados; se ve inmediatamente que el plano complejo **C** descrito por $m(z)$ y la esfera **S** son inversas en la inversión I(P, 2).

Esta aplicación $\mathbf{C} \to \mathbf{S}$ inyectiva según las fórmulas (1′) no es suprayectiva, pues $P(0, 0, -1)$ no es la imagen de ningún punto de **C**; pero si ponemos

$$z \in \mathbf{C} \qquad f(z) = M(X, Y, Z) \qquad f(\infty) = P(0, 0, -1)$$

definimos una *biyección* f de $\tilde{C}$ sobre **S** de la que la restricción **C** es la inversión **I(P, 2)**.

Pongamos ahora

$$(2) \qquad (\forall z_1, z_2 \in \tilde{C}) \qquad \delta(z_1, z_2) = D(M_1, M_2)$$

(esta fórmula comprende como caso particular $(z \in \mathbf{C})$ $\delta(z, \infty) = D(M, P)$); siendo D la distancia euclídea de $\mathbf{R}^3$, definimos así una distancia δ sobre $\tilde{C}$; llamaremos plano analítico(26) el *espacio métrico* así definido. La definición (2) muestra que es *homeomorfo* a la esfera **S** llamada *esfera* de RIEMANN asociada al plano analítico. El homeomorfismo f es igualmente una isometría, pues la distancia de z_1 y de z_2 (en $\tilde{C}$) es, por definición, igual a la distancia de M_1 y M_2 (en **S**).

(25) Se llama homeomorfismo de una parte A de un espacio métrico E sobre una parte **B** de un espacio métrico F toda biyección f de A sobre B tal que f y f^{-1} sean continuas para las distancias definidas, respectivamente, sobre E y F.

(26) En el libro de Geometría veremos que $\tilde{C}$ puede identificarse con el espacio proyectivo complejo de una dimensión, es decir, con la recta proyectiva compleja. (Se podría definir análogamente la recta proyectiva real $\tilde{\mathbf{R}}$ que contiene un *solo* punto en el infinito, que no debe confundirse con la recta numérica cerrada $\bar{\mathbf{R}}$ que contiene *dos* puntos en el infinito $-\infty$ y $+\infty$. Ver libro de Análisis.)

EJERCICIOS

1. Utilizando la definición de la distancia euclídea en $\mathbf{R}^3$,

$$D(M_1, M_2) = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2 + (Z_2 - Z_1)^2}$$

y una fórmula bien conocida de la inversión, demostrar que

$$(z_1 \in \mathbf{C},\ z_2 \in \mathbf{C}) \qquad \delta(z_1, z_2) = \frac{2\,|z_2 - z_1|}{\sqrt{(1 + |z_1|^2)(1 + |z_2|^2)}},$$

demostrar, por otra parte, que

$$(z \in \mathbf{C}) \qquad \delta(z, \infty) = \delta(\infty, z) = \frac{2}{\sqrt{1 + |z|^2}}.$$

Deducir que las cuatro biyecciones siguientes de $\tilde{\mathbf{C}}$ sobre sí mismo ($h \in \mathbf{C}$, $k \in \mathbf{C}^*$) son continuas, como también sus aplicaciones recíprocas (es decir, que son homeomorfismos de $\tilde{\mathbf{C}}$)

$$z \to z + h, \qquad z \to kz, \qquad z \to z^{-1}, \qquad z \to \bar{z}$$

(para el estudio de la continuidad a distancia finita se observará que si $|z_1| < r$, $|z_2| < r$ se tiene

$$\frac{2\,|z_2 - z_1|}{1 + r^2} < \delta(z_1, z_2) < 2\,|z_2 - z_1|$$

y al infinito se utilizará $\delta(\infty, z)$).

2. Demostrar que la restricción a $\mathbf{C}^*$ de la biyección de $\tilde{\mathbf{C}}$ sobre sí mismo definida por $z \to (\bar{z})^{-1}$ es la inversión I(O, 1) y que la restricción a $\mathbf{C}^*$ de la biyección de $\tilde{\mathbf{C}}$ sobre sí mismo definida sobre $z \to z^{-1}$ es el «producto» conmutativo de la inversión I(O, 1) y de la simetría respecto a Ox. Interpretar geométricamente ($k \in \mathbf{C}^*$)

$$z' = k/z, \qquad z' - z_0 = k/(z - z_0).$$

(Se supondrá primero k real, luego k no real.)

b) Función homográfica

Consideremos la aplicación f de $\tilde{\mathbf{C}}$ en sí mismo definida por ($a, b, c, d \in \mathbf{C}$)

$$f(z) = \frac{az + b}{cz + d} \qquad \text{si} \qquad z \neq \infty \qquad \text{y} \qquad z \neq -d/c$$

$$f(\infty) = a/c, \qquad f(-d/c) = \infty.$$

La restricción de f en $\mathbf{C}$ define una semejanza o una traslación si $c = 0$. Si $c \neq 0$

$$f(z) = \frac{a}{c} - \frac{ad - bc}{c^2} \, \frac{1}{z + \dfrac{d}{c}}$$

Si $ad - bc = 0$, f es constante, diremos que f es una *función homográfica impropia.* Si $ad - bc \neq 0$ (siendo c nulo o no) f es biyectivo, diremos que f es una *función homográfica propia.* Está compuesta de un número finito de aplicaciones de uno de los tipos siguientes ($h \in \mathbf{C}$, $k \in \mathbf{C}^*$)

$$z \to f_1(z) = z + h, \qquad z \to f_2(z) = kz, \qquad z \to f_3(z) = z^{-1}.$$

Estas aplicaciones son biyecciones de $\tilde{\mathbf{C}}$, las restricciones de f_1 y f_2 a $\mathbf{C}$ representan, respectivamente, una traslación y una semejanza y la restricción de f_3 a $\mathbf{C}^*$ representa una inversión-simetría (ver ej. 2 anterior). Las aplicaciones $M(z) \to M'(z')$ asociadas a las restricciones de $z \to z' = f_h(z)$ ($h = 1, 2, 3$) conservan los ángulos del plano orientado, y el conjunto de círculos y de rectas del plano se conservan globalmente por cada una de ellas. *Por abuso de lenguaje* diremos que f_1, f_2, f_3 y f, función homográfica, operando en $\tilde{\mathbf{C}}$, conservan los ángulos y el conjunto de rectas-círculos, de donde el nombre de *transformación circular* dada también a la función homográfica.

EJERCICIOS

3. Demostrar que toda función homográfica propia es homeomorfismo de $\tilde{\mathbf{C}}$ sobre sí mismo. (Se observará que el elemento compuesto de dos homeomorfismos es un homeomorfismo; es suficiente, pues, demostrar que f_1, f_2, f_3 son homeomorfismos) (ver ej. 1 anterior).

4. Si A y C son números reales y B complejo, demostrar que el conjunto de los puntos $M(z)$ verificando

$$(1) \qquad Az\bar{z} + Bz + \bar{B}\bar{z} + C = 0$$

es un círculo si $A \neq 0$ y una recta si $A = 0$. Demostrar que la ecuación (1) representa una recta si y sólo si (1) se verifica para $z = \infty$.

5. Se llama *puntos focales* de la función homográfica propia $z \to f(z) = (az + b)/(cz + d)$, los puntos de afijos $\varphi = -d/c$ (puntos focal objeto) y $\psi' = a/c$ (punto focal imagen), se tiene $f(\varphi) = \infty$ y $f(\infty) = \psi'$. ¿En qué condición la transformación homográfica transforma un círculo en círculo, un círculo en recta, una recta en círculo, una recta en recta?

Demostrar que la relación $z' = f(z)$ puede escribirse

$$(z - \varphi)(z' - \psi') = \lambda$$

hacer explícito el valor de λ.

c) Grupo circular

Si se pone

$$f(z) = \frac{az + b}{cz + d}, \qquad f'(z) = \frac{a'z + b'}{c'z + d'}$$

con $ad - bc \neq 0$ y $a'd' - b'c' \neq 0$ tendremos

$$(f' \circ f)(z) = \frac{(a'a + b'c)z + a'b + b'd}{(c'a + d'c)z + (c'b + d'd)}$$

con

$$(a'a + b'c)(c'b + d'd) - (a'b + b'd)(c'a + d'c) = (ad - bc)(a'd' - b'c')$$

en consecuencia, $f' \circ f$ es también una función homográfica propia (se sabía ya que era una biyección como compuesta de dos biyecciones). Por otro lado, $z \to z$ es una función homográfica propia, igualmente que

$$z \to f^{-1}(z) = \frac{-dz + b}{cz - a}$$

luego: *el conjunto de las transformaciones homográficas propias es un grupo que opera en el plano analítico. Se le llama el grupo circular* (sus elementos conservan globalmente el conjunto de círculos y rectas).

d) Razón doble de cuatro números complejos

En la traslación $z \to z' = z + h$ para todo par de puntos distintos z_1, z_2

$$z'_2 - z'_1 = z_2 - z_1$$

se dice que $z_2 - z_1$ es un *invariante para toda traslación.*

En la semejanza $z \to z' = kz$ $(k \neq 0)$, $z'_2 - z'_1 = k(z_2 - z_1)$, pero si consideramos tres puntos distintos

$$\frac{z'_1 - z'_3}{z'_2 - z'_3} = \frac{z_1 - z_3}{z_2 - z_3}$$

diremos que $(z_1 - z_3)/(z_2 - z_3)$ es un *invariante para toda semejanza y toda traslación.*

En fin, la transformación $z \to z' = z^{-1}$ tenemos

$$\frac{z'_1 - z'_3}{z'_2 - z'_3} = \frac{z_2}{z_1} \frac{z_1 - z_3}{z_2 - z_3}$$

pero si consideramos cuatro puntos distintos

$$\frac{z'_1 - z'_3}{z'_2 - z'_3} : \frac{z'_1 - z'_4}{z'_2 - z'_4} = \frac{z_1 - z_3}{z_2 - z_3} : \frac{z_1 - z_4}{z_2 - z_4} = \rho$$

este número ρ es un invariante para $z \to z^{-1}$, $z \to kz$ y $z \to z + h$, de donde:

TEOREMA Y DEFINICIÓN. — *Dados cuatro elementos de $\tilde{\mathbf{C}}$ distintos o no z_1, z_2, z_3, z_4, cuando el elemento de $\tilde{\mathbf{C}}$*

$$\rho = \frac{z_1 - z_3}{z_2 - z_3} : \frac{z_1 - z_4}{z_2 - z_4}$$

está definido, es invariante por toda función homográfica; se le llama la razón doble *de los cuatro números y se escribe $\rho = (z_1, z_2, z_3, z_4)$.*

Así, si $z_1 = z_3$ (z_1, z_2, z_4 distintos 2 a 2)

$$\rho = 0,$$

si $z_2 = z_3$ (z_1, z_2, z_4 distintos 2 a 2)

$$\rho = \infty,$$

si $z_4 = \infty$ (z_1, z_2, z_3 finitos, distintos 2 a 2)

$$\rho = \frac{z_1 - z_3}{z_2 - z_3}.$$

Corolario. — *Dados seis números complejos finitos o no z_1, z_2, z_3 los tres distintos y z'_1, z'_2, z'_3 los tres distintos, existe una función homográfica propia única f tal que $f(z_k) = z'_k$ $(k = 1, 2, 3)$.*

En efecto, si $z' = f(z)$ existe, se tendrá

$$(z', z'_1, z'_2, z'_3) = (z, z_1, z_2, z_3)$$

esta relación resuelta en z' muestra que f es una homografía propia y que $f(z_k) = z'_k$ $(k = 1, 2, 3)$.

EJERCICIOS

6. Demostrar que si se pone $(z_1, z_2, z_3, z_4) = \rho$ se tiene

$$(z_2, z_1, z_4, z_3) = (z_3, z_4, z_1, z_2) = (z_4, z_3, z_2, z_1) = \rho.$$

Deducir de lo anterior que si p es una permutación cualquiera de $\{1, 2, 3, 4\}$, luego un elemento de $\mathcal{S}_4$ (§ 85): $p(\rho) = (z_{p(1)}, z_{p(2)}, z_{p(3)}, z_{p(4)})$ toma al menos 6 valores. Demostrar que

$$(z_1, z_2, z_4, z_3) = 1/\rho \qquad (z_1, z_3, z_2, z_4) = 1 - \rho.$$

Deducir que los 6 valores que puede tomar $p(\rho)$ son ρ, $1/\rho$, $1 - \rho$, $1/(1 - \rho)$, $(\rho - 1)/\rho$, $\rho/(\rho - 1)$. Demostrar que las aplicaciones que hacen corresponder a ρ uno de los 6 valores precedentes describen un grupo isomorfo a $\mathcal{S}_3$ (ver capítulo 4, ej. 62).

7. Demostrar que $p(\rho)$ toma menos de 6 valores únicamente en los casos siguientes:

a) Dos números y sólo dos son iguales

$$p(\rho) \in \{0, \infty, 1\}.$$

b) $\rho = -1, \frac{1}{2}$ o 2,

$$p(\rho) \in \left\{ -1, \frac{1}{2}, 2 \right\}$$

cuando $\rho = -1$ se dice que la razón doble es *armónica.*

c) $\rho = -j$ o $-j^2$ $(j^3 = 1)$

$$p(\rho) \in \{-j, -j^2\}.$$

8. Dados cuatro puntos M_1, M_2, M_3, M_4 del plano euclídeo, de afijos z_1, z_2, z_3, z_4 (números complejos finitos), demostrar que la razón doble (z_1, z_2, z_3, z_4) es invariante

por un desplazamiento de ejes; se dirá que esta razón doble es la razón doble (M_1, M_2, M_3, M_4) de los 4 puntos; calcularlo en función de las longitudes de los vectores $\overrightarrow{M_iM_j}$ y de ciertos ángulos que estos vectores forman entre sí (utilizar el § 123, *d*). ¿Qué se puede decir de los 4 puntos si ρ es real, si ρ es real estrictamente positivo o estrictamente negativo? (utilizar el § 123, *d*).

e) Puntos invariantes de la función homográfica

Las soluciones α y β de $z = f(z)$ están dadas por

$$cz^2 + (d - a)z - b = 0.$$

Si $c \neq 0$ tendremos

$$\begin{array}{lll} \Delta = (d-a)^2 + 4bc = (a+d)^2 - 4(ad-bc) \neq 0 & \alpha \neq \beta & (\alpha \text{ y } \beta \text{ finitos}) \\ \Delta = 0 \quad y \quad d \neq a & \alpha = \beta & (\alpha \text{ finito}). \end{array}$$

En el caso en que $\Delta \neq 0$ para $z_1 \neq z_2$ tendremos

$$(z_1', z_2', \alpha, \beta) = (z_1, z_2, \alpha, \beta)$$

un cálculo fácil muestra que

$$\frac{z_1' - \alpha}{z_1' - \beta} : \frac{z_1 - \alpha}{z_1 - \beta} = \frac{z_2' - \alpha}{z_2' - \beta} : \frac{z_2 - \alpha}{z_2 - \beta}$$

si f es biyectiva existe luego un número complejo $s \neq 0$ tal que para todo z $(z' = f(z))$ se tiene

(1) $$\frac{z' - \alpha}{z' - \beta} = s\,\frac{z - \alpha}{z - \beta}$$

es la *forma canónica* de la función homográfica relativa a los puntos invariantes supuestos distintos.

En el caso $\Delta = 0$, $a \neq d$, un cálculo fácil nos permite encontrar la forma canónica $(h \in \mathbf{C})$

(2) $$\frac{1}{z' - \alpha} = \frac{1}{z - \alpha} + h.$$

EJERCICIOS

9. ¿Cuáles son los puntos invariantes (en $\tilde{\mathbf{C}}$) de una semejanza, de una traslación? Enunciar y demostrar los recíprocos.

Siendo f una función homográfica que tiene dos puntos invariantes distintos en $\mathbf{C}$, α y β, se designa por g la función $z \to g(z) = (z - \alpha)/(z - \beta)$, demostrar sin cálculo que $g \circ f \circ g^{-1}$ es una semejanza $z \to kz$. Calcular k.

10. Siendo s el número definido por la relación (1), demostrar que $s + \dfrac{1}{s}$ se expresa racionalmente en función de a, b, c, d.

11. Demostrar que los puntos focales (ej. 5 más arriba) y los puntos invariantes de una transformación circular tienen el mismo punto medio.

f) Transformación involutiva

Se dice que la función homográfica propia f es involutiva si $f = f^{-1}$ (o la condición equivalente: $f^2 = f \circ f$ es la identidad; § 15, ej. 6), ello es así si y sólo si $a + d = 0$; como

$$z' = f(z) \Leftrightarrow z = f(z')$$

se puede decir que z y z' son *homólogos* en la involución f.

EJERCICIOS

12. Demostrar que las únicas semejanzas involutivas son las simetrías con relación a un punto. Mostrar que las involuciones no reducidas a una semejanza tienen sus puntos invariantes α y β distintos y finitos y que

$$(z, z', \alpha, \beta) = -1.$$

Deducir que una involución está determinada por dos parejas (z_1, z'_1), (z_2, z'_2) de puntos homólogos.

Demostrar que los puntos focales (ej. 5) están confundidos y que la relación $z' = f(z)$ puede escribirse $(z - \varphi)(z' - \varphi) = \lambda$.

Interpretar geométricamente esta relación (utilizar el ejercicio 2 de este párrafo).

13. Demostrar que dos involuciones propias distintas de la identidad y distintas entre sí f y g tiene un par común y sólo uno (z_0, z'_0) de puntos homólogos. ¿Cuáles son los puntos dobles de la homografía $h = f \circ g$? Demostrar recíprocamente que toda homografía propia puede escribirse en la forma $f \circ g$, donde f y g son dos involuciones y esto de una infinidad de maneras.

Ejercicios

122. Sea **K** un cuerpo conmutativo totalmente ordenado, 1 su elemento unidad.

a) Demostrar que -1 no es un cuadrado de **K**

b) Demostrar que existe un supercuerpo **L** de **K**, conteniendo un subcuerpo **K'**, isomorfo a **K** (y que se identificará a **K**) y una raíz cuadrada de -1.

c) ¿Qué condición suplementaria debe verificar K para que todo elemento de L sea un cuadrado de L?

123. Calcular los números complejos siguientes

a) $\dfrac{(1+i)^4}{(1-i)^3} + \dfrac{(1-i)^4}{(1+i)^3}$.

b) $(1+i)a + (1-i)b$ con $a^2 = j - 1,\ b^2 = j^2 - 1$

(encontrar todos los valores posibles).

c) $z = \dfrac{f(x) - f(0)}{f(x) - f(\infty)}$ con $f(x) = \dfrac{1 + ix}{1 + i \operatorname{tg} a + ix(1 + i \operatorname{tg} b)}$

(a, b, x reales). Módulo y argumento de z.

124. Factorizar el polinomio: $x^2 + y^2 + z^2 - yz - zx - xy$ (se le considerará como un trinomio en x, y se introducirá j y j^2).

125. Demostrar que cualquiera que sean z y z' ($zz' = u^2$).

a) $|z + z'|^2 + |z - z'|^2 = 2(|z|^2 + |z'|^2)$.

Interpretación geométrica.

b) $|z| + |z'| = \left|\dfrac{z + z'}{2} - u\right| + \left|\dfrac{z + z'}{2} + u\right|$.

(si se demuestra que la relación precedente está conservada por similitud, se podrá suponer $u = 1$). Interpretación geométrica.

126. Encontrar todos los pares (z, z') de números complejos tales que

$$\sqrt{|zz'|} = \left|\frac{z + z'}{2}\right|$$

(la misma observación que en el ej. 125, **b**). Discutir

127. Calcular $\cos 5a$, $\operatorname{sen} 5a$, en función, respectivamente, de $\cos a$ y $\operatorname{sen} a$.

De $\cos 5\dfrac{\pi}{10} = 0$, deducir el valor de $\cos \dfrac{\pi}{10}$.

128. Se designa por α y $\bar{\alpha}$ las raíces de $x^2 - 2x + 2 = 0$.

a) Escribir α^n y $\bar{\alpha}^n$ (n entero positivo) en forma trigonométrica y en forma algebraica. Calcular

$$\prod_{k=0}^{n} (\alpha^k + \bar{\alpha}^k).$$

b) f y g son dos funciones reales de variable real derivables, se pone

$$h(t) = f(t) + ig(t), \quad h'(t) = f'(t) + ig'(t)$$

Calcular f y g para $h(t) = e^{(a+ib)t}$ (a, b reales); deducir que $h'(t) = (a + ib)\,h(t)$. Calcular la derivada n-ésima de $e^{\alpha t}$.

129. Dados los enteros n, p, k $(0 \leq k < p < n)$, se pone

$$S_k = C_n^k + C_n^{k+p} + C_n^{k+2p} + \ldots + C_n^x$$

siendo x el mayor entero posible.

a) Escribir x con ayuda de una parte entera (cap. 5, ej. 121).

b) Para $p = 3$, calcular S_0, S_1, S_2.

c) Para $p = 4$, calcular S_0, S_1, S_2, S_3 (ver § 92, ej. 4, utilizar $(1 + x)^n$ en donde x se ha elegido convenientemente.

130. Sea ABC un triángulo, los afijos de A, B, C son, respectivamente, a, b, c.

a) Calcular el afijo del centro de gravedad G de ABC.

b) Encontrar los puntos M, del afijo z, tales que

$$\overrightarrow{MA}/\|\overrightarrow{MA}\|^2 + \overrightarrow{MB}/\|\overrightarrow{MB}\|^2 + \overrightarrow{MC}/\|\overrightarrow{MC}\|^2 = 0$$

(se demostrará que $1/(z - a) + 1/(z - b) + 1/(z - c) = 0$).

131. Resolver la ecuación $(x - a)^n = k(x - a')^n$ (n entero positivo a, a', k complejos $k \neq 0$). Demostrar que estas raíces tienen sus imágenes sobre un círculo, o una recta. ¿En qué condiciones estas raíces son todas reales?

Resolver $\left(\dfrac{1 - ix}{1 + ix}\right)^n = \dfrac{1 - i \operatorname{tg} a}{1 + i \operatorname{tg} a}$ $\left(a \text{ real comprendido estrictamente entre } -\dfrac{\pi}{2} \text{ y } \dfrac{\pi}{2}\right)$

132. En el plano complejo se consideran los puntos $A_k(0 \leq k < n)$ de afijo

$$z_k = a\,(\cos 2k\pi/n + i \operatorname{sen} 2k\pi/n)\ (a > 0)$$

y el punto M de afijo $z = r\,(\cos\theta + i \operatorname{sen}\theta)$.

a) Demostrar que $z^n - a^n = (z - a)\,(z - z_1)\,\ldots\,(z - z_{n-1})$.

b) De la relación encontrada para $a = 1$, deducir

$$\operatorname{sen}\frac{\pi}{n}\operatorname{sen}\frac{2\pi}{n}\ldots\operatorname{sen}\frac{(n-1)\pi}{n} = n2^{1-n}$$

¿en qué se transforma esta fórmula para $n = 2p$ (p entero)?; aplicación numérica $n = 90$, $n = 100$.

c) Suponiendo $a \neq 1$, calcular el producto de las longitudes de los vectores $\overrightarrow{MA_k}$ $(0 \leq k < n)$ el producto de las longitudes de los vectores $\overrightarrow{A_0A_k}$ $(1 \leq k < n$ y el producto de las diagonales del polígono regular A_0, A_1, ..., A_{n-1}. (Tomar $0 \to 0$ y pasar al límite).

133. **a)** Conociendo los afijos ω y z de Ω y M, determinar el afijo de M′ imagen de M en la semejanza (Ω, α, k).

b) Conociendo los afijos a y b de dos puntos A y B, encontrar el afijo del vértice C del triángulo equilátero ABC de sentido directo.

c) Conociendo los afijos a_0, a_1 de dos vértices $A_0 A_1$ de un polígono regular de sentido directo $A_0 A_1 \ldots A_{n-1}$, calcular el afijo a_k de A_k.

d) Dado un rectángulo ABCD de sentido directo, calcular los afijos de C y D, conociendo los de A y B y la relación $k = BC/AB > 0$.

134. En el plano complejo, se consideran los cuatro puntos A_1, A_2, M_1, M_2 de afijos respectivos a, $-a$, z_1, z_2 (a real no nulo) verificando la relación $z_1 z_2 = a^2$.

a) Demostrar que Ox es bisectriz interior de

$$(\overrightarrow{OM_1}, \overrightarrow{OM_2}) \qquad \text{y que} \qquad \| \overrightarrow{OM_1} \| \, \| \overrightarrow{OM_2} \| = \| \overrightarrow{OA_1} \|^2.$$

b) Calcular la razón doble $(a, -a, z_1, z_2)$ deducir que las parejas (A_1, A_2) y (M_1, M_2) hacen el mismo papel.
Demostrar que estos cuatro puntos están o alineados o situados en un círculo C; en este último caso demostrar que las cuerdas A_1A_2 y M_1M_2 se cortan en el interior de C y son conjugadas en relación a C: se dice que el cuadrángulo $A_1A_2\ M_1M_2$ es *armónico.*

135. En el plano analítico, se considera la aplicación f que a $m(z)$ hace corresponder el punto M(Z) tal que

$$Z = f(z) = \frac{1}{2}\left(z + \frac{c^2}{z} \right)$$

(c real estrictamente positivo tiene por imagen F y $-c$, F').

a) Demostrar que M es la imagen por f de dos puntos m_1, m_2 y que F, F', m_1, m_2 forman o una división armónica o un cuadrángulo armónico (V. ej. 134).

b) Se pone $z' = g(z) = (z - c)/(z + c)$, calcular $f' = g \circ f \circ g^{-1}$. Deducir los conjuntos descritos por M (resp. m) cuando m (resp. M) describe

α) un círculo pasando por F y F'.

β) un círculo descrito por P tal que $PF/PF' = k$ ($k > 0$, $k \neq 1$).

c) Se pone $z = r$ (cos $\theta + i$ sen θ), $Z = X + iY$.
Calcular X, Y (reales) en función de r, θ, c.
¿Qué conjunto E (resp. H) describe M cuando m describe el círculo $r = R > 0$? (resp. la recta $\theta \equiv \alpha$ (mod π).
¿Qué conjunto describe m cuando M describe E, y cuando M describe H?

136. Tenemos dos números complejos a y b, se designa por z_1 y z_2 las raíces de las ecuación $x^2 - 2ax + b = 0$. ¿Qué condiciones deben satisfacer a y b para que z_1 y z_2 posean una de las propiedades siguientes?

1) $|z_1| = |z_2|$ 2) $\arg z_1 \equiv \arg z_2 \pmod{2\pi}$
3) $|z_1| = r\,|z_2|$ 4) $\arg z_1 \equiv \arg z_2 + \alpha \pmod{2\pi}$

(α y r son dos reales dados, $r > 0$).
Para 3) y 4), se pondrá $z = z_1/z_2$ y calculando $z + 1/z$, se utilizará el ej. 135, *c*).

137. Se considera la transformación circular f definida por (V. § 124)

$$z \to z' = f(z) = \frac{az + b}{cz + d}$$

con $c(ad - bc) \neq 0$; se designará por la misma letra todo punto y su afijo. Sean u y v los puntos invariantes de f, se pondrá

$$\text{si } u \neq v, \; Z = g(z) = \frac{z - u}{z - v}$$

$$\text{si } u = v, \; Z = g(z) = \frac{1}{z - u}$$

y en los dos casos $F = g \circ f \circ g^{-1}$.

a) Si $u \neq v$, demostrar que existe un número complejo $s \neq 0$, tal que $F(Z) = sZ$. Deducir de ello que f conserva globalmente el haz $\mathfrak{F}$ de los circuitos que pasan por u y v y el haz conjugado $\mathfrak{F}'$. Demostrar que para que f conserve *cada uno* de los circuitos (o rectas) de $\mathfrak{F}$ (resp $\mathfrak{F}'$) es necesario y suficiente que s sea real (resp. $|s| = 1$). ¿Qué ocurre para $s = -1$?

b) Si $u = v$, demostrar que existe un número complejo h tal que $F(Z) = Z + h$. Demostrar que existe un haz de círculos $\mathfrak{F}$ en que cada círculo es invariante por f y un haz de círculos $\mathfrak{F}'$ globalmente invariante por f.

c) Se pone $z_1 = f(z_0)$ ... $z_n = f(z_{n-1})$. Calcular z_n. ¿En qué caso para todo z, se tiene $z_n = z_0$?
¿Cómo están entonces dispuestos los puntos $z_0, \ldots, z_{n-1}$?
Caracterizar f si este fenómeno se produce para $n = 2$.

138. Se considera el grupo G de transformaciones del plano analítico $\tilde{C}$ engendrado por f_1 y f_2 definidas por $f_1(z) = 1/z$, $f_2(z) = 1 - z$.

a) Demostrar que G tiene seis elementos (V. cap. 4, ej. 62) que se designarán por f_0 con $f_0(z) = z$, f_1, f_2, f_3, f_4, f_5.

b) Siendo C_0 el círculo $|z| = 1$, demostrar que $C_k = f_k(C_0)$ es un círculo o una recta. Determinar para cada valor de k la imagen por f_k de $|z| \leq 1$ y $|z| \geq 1$ $(0 \leq k \leq 5)$.

c) Demostrar que las curvas C_k encontradas parten el plano en seis regiones Δ_h $(0 \leq h \leq 5)$ y que si $z = f_0(z)$ está contenido en una de ellas estrictamente, cada una de las cinco restantes contienen estrictamente uno y sólo un punto $f_k(z)$ $(1 \leq k \leq 5)$.

d) ¿Qué relación hay entre este ejercicio y los ejercicios 6 y 7 del § 124?

139*. Se considera el conjunto F de las funciones homográficas propias f operando en el plano analítico $\tilde{\mathbf{C}}$ y definidas por

$$f(z) = \frac{az + b}{cz + d} \qquad \text{si } z \neq \infty \qquad \text{y} \qquad z \neq -\frac{d}{c}$$

$$f(\infty) = \frac{a}{c} \qquad f\left(-\frac{d}{c}\right) = \infty \qquad ad - bc \neq 0.$$

Poniendo $z = x + iy$ $(x, y \in \mathbf{R})$ se designa por P (resp. P_1) el subconjunto de $\mathbf{C}$ definido por $y > 0$ (resp. $y < 0$), por D el subconjunto $|z| = 1$ y Γ el subconjunto $z = 1$. Finalmente se designa por G, G_1, H los subconjuntos de F descritos, respectivamente, por g, g_1, h tales que

$$g(\mathbf{P}) = \mathbf{P}, \quad g_1(\mathbf{P}) = \mathbf{P}_1, \quad h(\mathbf{D}) = \mathbf{D}.$$

a) Demostrar que $g(\mathbf{P}_1) = \mathbf{P}_1$, $g(\mathbf{R}) = \mathbf{R}$, $g_1(\mathbf{P}_1) = \mathbf{P}$ y $g_1(\mathbf{R}) = \mathbf{R}$. Demostrar recíprocamente que $f(\mathbf{R}) = \mathbf{R}$ implica $f \in G$, o sea, $f \in G_1$. (Utilizar la continuidad de f, § 124, ej. 3.) Demostrar que los elementos de G pueden estar definidos por a, b, c, d reales verificando $ad - bc = 1$ (se observará que si $\lambda \neq 0$ (a, b, c, d) y $(\lambda a, \lambda b, \lambda c, \lambda d)$ definen la misma función homográfica).

b) Demostrar que G es un subgrupo del grupo circular F; ¿es transitivo este grupo?

¿G_1 es un subgrupo de F?

c) Encontrar una aplicación f_0 tal que $f_0(P) = D$, $f_0(R) = \Gamma$ (utilizar ej. 2 § 116). Demostrar que si g describe G, entonces $h = f_0 \circ g \circ f_0^{-1}$ describe H. Deducir que $h(\Gamma) = \Gamma$ y que H es un subgrupo del grupo F.

Demostrar que h puede estar definida por

$$h(z) = \omega \frac{z - z_0}{1 - \bar{z}_0 z} \qquad |\omega| = 1 \qquad |z_0| < 1.$$

140. Si A, B, C, D son cuatro números complejos tales que $AD - BC \neq 0$, se considerará la aplicación f de $\tilde{C}$ (plano analítico) en sí mismo definida por

$$(1) \qquad z' = f(z) = \frac{A\bar{z} + B}{C\bar{z} + D} \quad \text{si} \quad z \neq \infty \quad \text{y} \quad z \neq -\frac{\bar{D}}{\bar{C}}$$

$$f(\infty) = \frac{A}{C}, \qquad f\left(-\frac{\bar{D}}{\bar{C}}\right) = \infty$$

f se llama una *antihomografía*.

1. a) Demostrar que f es una biyección y que un cambio de ejes rectangulares de la misma orientación no altera la forma de la relación (1).

b) ¿Cómo transforma f la razón doble de cuatro puntos? ¿Se puede conservar este último?

Demostrar que f conserva el conjunto de las rectas-círculos.

2. Se supone $C = 0$.

a) Demostrar que si se escoge convenientemente los ejes, f puede estar definido por

$$(2) \qquad z' = r\bar{z} + p + iq$$

donde p, q, r son reales y $r > 0$.

b) Interpretar geométricamente (2). (Discutir buscando los puntos intervariantes de f).

3. Se supone $C \neq 0$ y f involutivo (es decir, $f = f^{-1}$), se dice que f es una antiinvolución.

a) Demostrar que f está definida por

$$(3) \qquad z'\bar{z} - (p - iq)z' - (p + iq)\bar{z} + r = 0$$

donde p, q, r son reales.

b) Interpretar geométricamente (3). (Buscar los puntos invariantes de f y efectuar una traslación de ejes.)

4. Se supone $C \neq 0$ y $f \neq f^{-1}$.

a) Demostrar que $f^2 = f \circ f$ es una homografía. Deducir sin cálculo que las funciones f pertenecen a uno de los tres tipos siguientes

(T_1) hay dos puntos $u \neq v$ y dos sólo tales que

$$f(u) = v \qquad f^{-1}(u) = v$$

y u y v son finitos.

(T_2) existen dos puntos $u \neq v$ y dos solo invariantes por f y u y v son finitos.

(T_3) existe un solo punto u invariante por f y u es finito.

b) Demostrar que mediante un cambio de ejes las funciones f del tipo (T_1) y del tipo (T_2) pueden ser definidas, respectivamente, por

(4) $$z'\bar{z} - \lambda(z' + \bar{z}) + a^2 = 0$$

(5) $$z'\bar{z} - \lambda(z' - \bar{z}) - a^2 = 0$$

λ es complejo y a real estrictamente positivo.

Se pone $Z = g(z) = \dfrac{z - a}{z + a}$, $\quad F = g \circ f \circ g^{-1}$.

Interpretar geométricamente las funciones F que corresponden, respectivamente, a las funciones f de los tipos (T_1) y (T_2).

¿Cómo f del tipo (T_1) (resp. del tipo T_2) transforma los círculos pasando por a y $-a$, o los círculos del haz conjugado? ¿Hay círculos o rectas invariantes por f del tipo (T_1) (resp. del tipo T_2)?

c) Demostrar que mediante un cambio de ejes las funciones f de tipo (T_3) pueden ser definidas por

$$\frac{1}{z'} = \frac{1}{\bar{z}} + \lambda$$

siendo λ complejo.

Se pone $Z = h(z) = \dfrac{1}{z}$, $\quad F = h \circ f \circ h^{-1}$.

Interpretar geométricamente F. ¿Existen círculos o rectas invariantes por f?

7

ESPACIOS VECTORIALES

I. Definiciones. Primeras propiedades

125. Estructura de espacio vectorial sobre un cuerpo conmutativo. Isomorfismo de espacios vectoriales. Ejemplos

a) DEFINICIÓN. — *Dado un cuerpo conmutativo* K, *de elementos neutros* 0 *y* e, *se dice que un conjunto* E *provisto de una operación interna y de una operación externa cuyo dominio de operadores es* K, *tiene una* estructura de espacio vectorial sobre K *si:*

— E *es un grupo abeliano para su operación interna.*

— *La operación externa es tal que, para todo* x *de* E *y todo* λ *y todo* μ *de* K, *se tenga*

$$\lambda(\mu x) = (\lambda\mu)x \qquad ex = x.$$

— *La operación externa es distributiva con relación a la suma en* K *y con relación a la operación interna de* E.

Se dice también que E es un espacio vectorial sobre K para las operaciones consideradas, o un espacio vectorial sobre K o también un espacio vectorial si no puede surgir confusión alguna. Designando multiplicativamente la ley externa de E y aditivamente su ley interna, 0_E y 0_K representan —provisionalmente— los elementos ceros de E y de K, los *axiomas* de la estructura de espacio vectorial sobre K se escriben

$$\begin{cases} V_1 & (\forall x, y, z \in K) & (x+y)+z = x+(y+z) \\ V_2 & (\exists 0_E \in E)(\forall x \in E) & 0_E + x = x + 0_E = x \\ V_3 & (\forall x \in E)(\exists x' \in E) & x + x' = x' + x = 0_E \qquad (x' = -x) \\ V_4 & (\forall x, y \in E) & x + y = y + x \end{cases}$$

$$\begin{cases} V_5 & (\forall x \in E)(\forall \lambda, \mu \in K) & \lambda(\mu x) = (\lambda\mu)x \\ V_6 & (\forall x \in E) & ex = x \end{cases}$$

$$\begin{cases} V_7 & (\forall x \in E)(\forall \lambda, \mu \in K) & (\lambda + \mu)x = \lambda x + \mu x \\ V_8 & (\forall x, y \in E)(\forall \lambda \in K) & \lambda(x+y) = \lambda x + \lambda y. \end{cases}$$

E para su adición interna es un grupo: se le llama *grupo aditivo* de E. x, y, z, ..., elementos de E se llaman los *vectores* de E, λ, μ, ..., elementos del cuerpo de operadores K se les llama *escalares*: los designaremos en general, respectivamente, por letras latinas y griegas, sin que este convenio sea absoluto, lo que nos llevaría a complicar las notaciones.

K se llama el *cuerpo de base* del espacio vectorial E; recordemos que le suponemos conmutativo sin que tengamos necesidad de repetirlo cada vez que hablemos de él.

Observemos que si K′ es un *subcuerpo* de K, la suma interna de E y la restricción de la ley externa a $E \times K'$ proporcionan a E una estructura de espacio vectorial sobre K′.

b) Isomorfismo de dos espacios vectoriales sobre el mismo cuerpo conmutativo K

DEFINICIÓN. — *Sean dos espacios vectoriales* E *y* E′ *sobre el mismo cuerpo* K, *toda biyección* f *de* E *sobre* E′ *tal que*

$$(1) \qquad \begin{array}{ll} (\forall x \in E)(\forall y \in E) & f(x+y) = f(x) + f(y) \\ (\forall x \in E)(\forall \alpha \in K) & f(\alpha x) = \alpha f(x) \end{array}$$

es un isomorfismo *de* E *sobre* E′.

Si $E = E'$, f *es un* automorfismo *del espacio vectorial* E.

Estudiaremos con detalle los isomorfismos de espacios vectoriales como caso particular de los homomorfismos en la sección III. Hasta entonces nos basta saber que la composición de dos isomorfismos es un isomorfismo y que f^{-1} es un isomorfismo de E′ sobre E, se comprobará fácilmente (lo demostraremos en el § 140). Por otro lado, una buena parte de las demostraciones se han hecho en el § 77: un isomorfismo del espacio vectorial E sobre el espacio vectorial E′ es según (1) un isomorfismo del grupo aditivo E sobre el grupo aditivo E′.

Si existe un isomorfismo f: $E \to E'$, se dice que E y E′ son dos *espacios vectoriales isomorfos*: recordemos que éstos son dos espacios vectoriales sobre el mismo cuerpo K; se dice entonces que E y E′ tienen la misma estructura de espacio vectorial.

EJEMPLOS Y EJERCICIOS

Se demostrará que los axiomas de la estructura de espacio vectorial sobre un cuerpo se verifican para los conjuntos E y los cuerpos K siguientes:

1. E: conjunto de los vectores libres del espacio (o del plano) de la geometría elemental, K = **R**, E está provisto de las operaciones $\vec{x} + \vec{y}$ y $\alpha\vec{x}$ (α real). Aquí está el origen de la palabra «espacio vectorial».

2. E: conjunto de polinomios con coeficientes reales, K = **R**, E posee dos operaciones A + B y αA (α real) (ver capítulo 11).

3. E es el mismo cuerpo K, las dos operaciones son naturalmente $a + b$ y ab, por tanto:

Todo cuerpo conmutativo es un espacio vectorial sobre sí mismo.

4. E = **C**, K = **R**, el conjunto **C** está provisto de dos operaciones

$$(a + a'i) + (b + b'i) = (a + b) + (a' + b')i \quad \text{y} \quad c(a + a'i) = ca + ca'i$$
(a, a', b, b', c son números reales).

C es un espacio vectorial sobre **R**, es también un espacio vectorial sobre sí mismo (ej. 3 anterior): estas dos estructuras de espacios vectoriales son diferentes (ver § 136, ej. 2).

E′ = **R** × **R** provisto de las dos operaciones $(a, a') + (b, b') = (a + b, a' + b')$, $c(a, a') = (ca, ca')$, (donde a, a', b, b', c son reales) es un espacio vectorial sobre **R** (ver § 127) isomorfo a **C** considerado como un espacio vectorial sobre **R**.

5. E = $\mathfrak{F}$(**R**, **R**) conjunto de funciones numéricas reales de variable real $t \to f(t)$, K = **R**, E posee dos operaciones (α real)

$$s = f + g \Leftrightarrow [(\forall t \in \mathbf{R}) \quad s(t) = f(t) + g(t)]$$
$$h = \alpha f \quad \Leftrightarrow [(\forall t \in \mathbf{R}) \quad h(t) = \alpha f(t)]$$

o F = $\mathfrak{F}$([0, 1], **R**), K = **R**, F está provisto de las operaciones anteriores.

6. E: conjunto de sucesiones (u_n) reales que verifican (a, b, reales)

$$u_{n+2} + au_{n+1} + bu_n = 0$$

K = **R**, E está provisto de las operaciones $u_n + v_n$, αu_n (α real).

7. E: conjunto de las funciones numéricas reales de variables reales dos veces derivables verificando (a, b reales)

$$f'' + af' + bf = 0$$

K = **R**, E está provisto de las operaciones definidas en el ejemplo 5 anterior.

8. Para todo α no nulo de K, $x \to \alpha x$ es un automorfismo del espacio vectorial **E** sobre K. Se le llama *homotecia* de razón $\alpha \neq 0$.

126. Reglas de cálculo en un espacio vectorial

a) La multiplicación externa es distributiva respecto a la resta en E; en efecto, para todo λ, todo x y todo y

$$\lambda(x - y) + \lambda y = \lambda[(x - y) + y] = \lambda x$$

de donde

$$\lambda(x - y) = \lambda x - \lambda y$$

si hacemos $x = y$ obtenemos

(1) $$\lambda 0_E = 0_E$$

finalmente

$$\lambda(-x) = \lambda(0_E - x) = \lambda 0_E - \lambda x = 0_E - \lambda x = -\lambda x.$$

b) La multiplicación externa es distributiva respecto a la resta en K; en efecto, para todo λ, todo μ y todo x

$$(\lambda - \mu)x + \mu x = [(\lambda - \mu) + \mu]x = \lambda x$$

de donde

$$(\lambda - \mu)x = \lambda x - \mu x$$

haciendo $\lambda = \mu$ obtenemos

(2) $$0_K x = 0_E$$

finalmente

$$(-\lambda)x = (0_K - \lambda)x = 0_K x - \lambda x = 0_E - \lambda x = -\lambda x.$$

c) Para todo λ de K y todo x de E, (1) y (2) demuestran que

(3) $$0_K x = \lambda 0_E = 0_E$$

consideremos recíprocamente la igualdad

$$\lambda x = 0_E$$

o bien $\lambda = 0_K$ (es el caso de (2)) o bien $\lambda \neq 0_K$ y λ es inversible en el cuerpo conmutativo K, de donde

$$\lambda^{-1}(\lambda x) = \lambda^{-1} 0_E = 0_E$$

y utilizando los axiomas V_5 y V_6

$$\lambda^{-1}(\lambda x) = (\lambda^{-1}\lambda)x = e \cdot x = x = 0_E.$$

En todo espacio vectorial la igualdad $\lambda x = 0_E$ *implica* $\lambda = 0_K$ *o* $x = 0_E$.

Este resultado y las igualdades (3) demuestran que *no hay, en general, ningún inconveniente en representar* 0_E *y* 0_K *por el mismo símbolo* 0, lo que haremos en adelante, salvo en el caso particular en que las notaciones 0_E y 0_K exijan una precisión determinada. Igualmente designaremos e elemento unidad de K por 1, lo que no tendrá ningún inconveniente, en general: los cuerpos que se utilizan normalmente son **Q**, **R**, **C** o cuerpos de característica nula.

127. Espacio vectorial producto de dos espacios vectoriales sobre el mismo cuerpo conmutativo K

Sean E_1 y E_2 dos espacios vectoriales sobre el mismo cuerpo conmutativo K, consideremos el grupo producto de los grupos aditivos E_1 y E_2, la operación sobre $E_1 \times E_2$ está definida por (ver § 79)

$$(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2).$$

Se verificará fácilmente que la multiplicación externa cuyo dominio de operadores es K, definido sobre $E_1 \times E_2$ por la igualdad

$$\lambda(x_1, x_2) = (\lambda x_1, \lambda x_2)$$

proporciona al grupo $E_1 \times E_2$ una estructura de espacio vectorial sobre K, se le llama *espacio vectorial producto* $E_1 \times E_2$.

Dados n espacios vectoriales $E_1, \ldots, E_n$ sobre K, las igualdades siguientes (con notaciones evidentes)

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$$
$$\lambda(x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n)$$

definen el espacio vectorial $E_1 \times E_2 \times \ldots \times E_n$. Se puede tomar $E_1 = E_2 = \ldots = E_n$. En particular, K es un espacio vectorial sobre K (§ 125, ej. 3); por tanto, K^n es un espacio vectorial sobre K, tendremos

$$a + b = (\alpha_1, \alpha_2, \ldots, \alpha_n) + (\beta_1, \beta_2, \ldots, \beta_n) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \ldots, \alpha_n + \beta_n)$$
$$\lambda a = \lambda(\alpha_1, \alpha_2, \ldots, \alpha_n) = (\lambda\alpha_1, \lambda\alpha_2, \ldots, \lambda\alpha_n)$$

por ejemplo, $\mathbf{R}^n$ es un espacio vectorial sobre **R**.

EJERCICIO

$\mathbf{C}^n$ es un espacio vectorial sobre **R**, es isomorfo a $\mathbf{R}^{2n}$ espacio vectorial sobre **C**.

II. Subespacios vectoriales

128. Definición. Ejemplos

a) TEOREMA Y DEFINICIÓN. — *Toda parte no vacía* F *de un espacio vectorial* E *sobre* K, *estable para la adición interna de* E *y la multiplicación externa, tiene por estructura inducida sobre ella por la estructura de* E *una estructura de espacio vectorial sobre* K. *Se dice que* F *es un* subespacio vectorial *de* E.

En efecto, si F es estable para la multiplicación externa y contiene x contiene $(-1)x = -x$ (§ 126, *b*); en consecuencia, el axioma V_3 se verifica en F. Por otra parte, todos los demás axiomas del § 125 se verifican en F. Daremos, pues, el resultado siguiente:

TEOREMA. — *Para que una parte no vacía* F *de un espacio vectorial* E *sobre* K, *sea un subespacio vectorial de* E, *es necesario y suficiente que*

(1) $(\forall x \in E) \ (\forall y \in F) \ x - y \in F$

(2) $(\forall x \in E) \ (\forall \alpha \in K) \ \alpha x \in F.$

OBSERVACION

(1) expresa el hecho de que el subespacio F del espacio vectorial E es un subgrupo del grupo aditivo E; según la demostración anterior, se puede reemplazar (1) (2) por las condiciones equivalentes

(1′) $(\forall x \in E) \ (\forall y \in F) \ x + y \in F$

(2) $(\forall x \in E) \ (\forall \alpha \in K) \ \alpha x \in F$

pero (1′) sólo no implica que F sea un subgrupo del grupo aditivo E.

Un subespacio vectorial de E *no es nunca vacío*: contiene siempre 0; el menor de todos (para la inclusión) es $\{0\}$, el mayor E; todo subespacio distinto de $\{0\}$ y E se llama subespacio vectorial *propio de* E.

Si x es un elemento determinado de E, la parte descrita por λx, cuando λ describe K es como se verificará sin dificultad, un subespacio vectorial de E, es distinto de $\{0\}$ si y sólo si $x \neq 0$.

b) Sean p elementos $x_1, x_2, \ldots, x_p$ de E, todo elemento de la forma

$$\lambda^1 x_1 + \lambda^2 x_2 + \ldots + \lambda^p x_p = \sum_{i=1}^{p} \lambda^i x_i$$

donde $\lambda^1, \lambda^2, \ldots, \lambda^p$ son escalares, se llama *combinación lineal* de $x_1, x_2, \ldots, x_p$.

ATENCIÓN: En λ^i, i es un *índice* y no un exponente; es por razones de comodidad, que se verán a continuación, que se adopta esta notación.

Se verificará fácilmente que el *conjunto de todas las combinaciones lineales, de p elementos determinados* $x_1, x_2, \ldots, x_p$ *de* E *es un subespacio vectorial de* E.

Más generalmente si A es una parte finita o infinita del espacio vectorial E sobre K, cuyos elementos están en correspondencia biyectiva con los de un conjunto de índices I, se llama *combinación lineal finita* de elementos de A todo elemento de E de la forma $\sum_{i \in I} \lambda^i x_i$, donde sólo un número finito de escalares λ^i son no nulos. Se verificará que el *conjunto de todas las combinaciones lineales finitas de elementos de* $A \subset E$ *es un subespacio vectorial de* E.

EJEMPLOS Y EJERCICIOS

1. En el espacio vectorial E sobre **R** de los vectores libres del espacio (§ 125, ej. 1), el conjunto de los vectores libres paralelos a una recta (resp. a un plano) es un subespacio de E.

2. El conjunto de los polinomios en x con coeficientes reales de grado inferior o igual a n (comprendido el polinomio cero) es un subespacio del espacio vectorial de los polinomios en x con coeficientes reales (§ 125, ej. 2, y capítulo 11).

3. **R** espacio vectorial sobre sí mismo es un subespacio de **C** espacio vectorial sobre **R** (§ 125, ej. 4).

4. En $E = \mathfrak{F}(\mathbf{R}, \mathbf{R})$ espacio vectorial sobre **R** (§ 125, ej. 5), determinar entre los subconjuntos siguientes aquellos que son subespacios de E:

a) Conjunto de las funciones pares.
b) Conjunto de las funciones impares.
c) Conjunto de las funciones continuas.
d) Conjunto de las funciones derivables.
e) Conjunto de las funciones k veces derivables.
f) Conjunto de las funciones definidas § 125, ej. 7.
g) Conjunto de las funciones tales que para todo t, $f(t_0 - t) = f(t)$ o bien $f(t_2) = \alpha f(t_1)$ o bien $f(t_2) = f(t_1) + \alpha$ o bien $f(t_0) = \alpha$ (t_0, $t_1 \neq t_2$, α números reales fijos).

5. Si E_1 y E_2 son dos espacios vectoriales sobre K, $E'_1 = E_1 \times \{0\}$ subconjunto de $E_1 \times E_2$ descrito por $(x_1, 0)$ es un subespacio vectorial de $E_1 \times E_2$; es isomorfo a E_1. Igualmente $E'_2 = \{0\} \times E_2$ es un subespacio de $E_1 \times E_2$ isomorfo a E_2.

Se identifica a menudo E'_1 y E_1 (y E'_2 y E_2), lo que permite decir que E_1 y E_2 son subespacios vectoriales de $E_1 \times E_2$.

129. Intersección de subespacios vectoriales. Subespacio vectorial engendrado por una parte A de un espacio vectorial E

a) La intersección de dos subespacios vectoriales F_1 y F_2 de un espacio vectorial E no es nunca vacía (contiene 0); por un lado, si x e y pertenecen a F_1 y F_2, lo mismo sucede con $x - y$ y con λx para todo λ de K; por tanto, $x - y$ y λx pertenecen a $F_1 \cap F_2$, que es un subespacio vectorial de E. De una manera más general, sea $\mathfrak{F}$ una familia cualquiera de subespacios vectoriales de E, consideremos su intersección (ver § 7)

$$I = \bigcap_{F \in \mathfrak{F}} F;$$

no es vacía (contiene 0) si x e y pertenecen a todo F de $\mathfrak{F}$, $x - y$ y λx (para todo λ de K) pertenecerán a I, por tanto:

Toda intersección de subespacios vectoriales de E *es un subespacio vectorial de* E.

b) Consideremos en particular una parte no vacía A de E, existen subespacios vectoriales de E que contienen A, por ejemplo, E. Consideremos la intersección de esta familia no vacía de subespacios vectoriales de E, es un subespacio vectorial de E y es el menor (para la inclusión de conjuntos), se le llama el *subespacio vectorial engendrado por* A.

Busquemos, por ejemplo, el subespacio vectorial F engendrado por $A = \{x_1, x_2, \ldots, x_p\}$, F contiene todas las combinaciones lineales de $x_1, x_2, \ldots, x_p$ que describen el subespacio vectorial F'; por tanto, $F' \subset F$.

Pero F′ contiene A y F es el menor subespacio vectorial conteniendo A; por tanto, F′ = F, de donde:

TEOREMA Y DEFINICIÓN. — *El menor subespacio vectorial* F *conteniendo p elementos de* E, $x_1, x_2, \ldots, x_p$, *es decir, el subespacio engendrado por* $\{x_1, x_2, \ldots, x_p\}$ *es el subespacio de las combinaciones lineales de* $x_1, x_2, \ldots, x_p$.

Se dice que $A = \{x_1, x_2, \ldots, x_p\}$ *es una* parte generatriz(27) *de* F; *cuando el mismo espacio vectorial* E *está engendrado por un número finito de sus elementos se dice que* E *es de* dimensión finita.

Veremos el origen de este término "dimensión finita" en el § 136; el grupo aditivo de un espacio vectorial de dimensión finita es de tipo finito (§ 82). Observemos que no es necesario suponer aquí los p vectores $x_1, x_2, \ldots, x_p$ distintos. Si se suprimen todos los que son iguales entre sí, menos uno, o los que son nulos, no se modifica el subespacio engendrado por el sistema considerado.

EJERCICIOS

1. A es una parte finita o infinita de un espacio vectorial E sobre K, demostrar que el subespacio engendrado por A es el subespacio de las combinaciones lineales finitas de elementos de A (§ 128, *b*).

2. En el espacio vectorial de los polinomios en x con coeficientes reales (o complejos), ¿cuál es el subespacio engendrado por x^2 y x^5? (§ 125, ej. 2).

3. En el espacio vectorial sobre **R** de las aplicaciones de **R** en **R** el subespacio engendrado por las funciones

$$t \rightarrow \text{sen } nt, \qquad t \rightarrow \cos nt$$

n describiendo **N** es el subespacio descrito por los «polinomios trigonométricos» (§ 121, *b*)

$$P(t) = a_0 + a_1 \cos t + b_1 \text{ sen } t + \ldots + a_n \cos nt + b_n \text{ sen } nt,$$

donde a_0, a_n, b_n son reales y n describe **N**.

130. Espacio vectorial cociente

Dado un espacio vectorial E sobre el cuerpo K, y una relación de equivalencia R definida sobre E, diremos que la relación R es *compatible con la estructura de espacio vectorial sobre* K *si*

(1) $[x \equiv y,\ x' \equiv y' \pmod{R}] \Rightarrow [x + x' \equiv y + y' \pmod{R}]$

(2) $(\forall \lambda \in K)[x \equiv y \pmod{R}] \Rightarrow [\lambda x \equiv \lambda y \pmod{R}]$

(27) Preferimos más la expresión *parte generatriz* que la expresión tradicional *sistema de generadores*, que podría conducir a creer que cada elemento de A es un generador de F, lo que es falso; es como conjunto que $\{x_1, x_2, \ldots, x_p\}$ engendra F. O bien sería preciso escribir «sistema — de — generadores», locución global en la que el «de» no tendría su valor de preposición. Diremos igualmente que (x_i) es una *familia generatriz* de F (ver § 82, nota (1)).

(1) expresa que R es compatible con la estructura del grupo aditivo de E, hemos visto (§ 75) que la relación R es entonces de la forma

$$x - y \in F \tag{3}$$

donde F es un subgrupo del grupo aditivo E.

(2) expresará entonces que si $x - y$ pertenecen a F, igualmente $\lambda(x - y)$, cuando λ es un elemento cualquiera de K; pero $x - y$ también es un elemento cualquiera de F; por tanto, (1) y (2) expresan que F es un subespacio vectorial de E. En el conjunto cociente E/R, designado E/F, de las clases $(\dot{x})$ módulo F, se podrá definir una adición interna y una multiplicación externa por las igualdades

$$(\dot{x}) + (\dot{y}) = \left(\dot{\widehat{x + y}}\right) \tag{4}$$

$$\lambda(\dot{x}) = \left(\dot{\widehat{\lambda x}}\right). \tag{5}$$

(4) da a E/F una estructura de grupo abeliano (§ 75), se comprobará fácilmente que (4) y (5) dan a E/F una estructura de espacio vectorial sobre K, de donde:

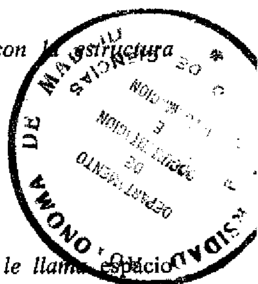
Teorema y definición. — *Toda relación de equivalencia compatible con la estructura de un espacio vectorial sobre* K *es de la forma*

$$x - y \in F$$

donde F *es un subespacio vectorial de* E. *Las igualdades*

$$\dot{x} + \dot{y} = \dot{\widehat{x + y}} \qquad \lambda\dot{x} = \dot{\widehat{\lambda x}}$$

proporcionan al conjunto cociente una estructura de espacio vectorial; se le llama espacio vectorial cociente de E *por* F *y se le designa* E/F.



EJERCICIO

Determinar los diferentes espacios cocientes relativos a los espacios vectoriales y a sus subespacios indicados en el § 128.

111. Suma de dos subespacios vectoriales. Subespacios suplementarios

a) Si E_1 y E_2 son dos subespacios vectoriales del espacio vectorial E sobre K, la parte de E descrita por $x_1 + x_2$, donde x_1 describe E_1 y x_2, E_2 es un subgrupo aditivo de E designado $E_1 + E_2$ (§ 79); además, cualquiera que sean λ de K, x_1 de E_1 y x_2 de E_2

$$\lambda(x_1 + x_2) = \lambda x_1 + \lambda x_2 \in E_1 + E_2$$

por tanto, $E_1 + E_2$ es un subespacio de E (ver § 127); por otra parte, es claramente el menor subespacio que contiene E_1 y E_2, en consecuencia:

Teorema 1 y definición. — *Si* E_1 *y* E_2 *son dos subespacios vectoriales de* E, *el conjunto* $E_1 + E_2$ *es el subespacio vectorial engendrado por* $E_1 \cup E_2$, *se le llama* subespacio suma *de* E_1 *y de* E_2.

OBSERVACION

Se verá que $E_1 \cup E_2$ *no es*, en general, *un subespacio vectorial de* E (ver ej. 4 a continuación).

b) Consideremos un espacio vectorial E y dos de sus subespacios E_1 y E_2 tales que $E = E_1 + E_2$, es decir,

$$(\forall x_1 \in E) \qquad (\exists x_1 \in E_1) \qquad (\exists x_2 \in E_2) \qquad x = x_1 + x_2$$

en general esta descomposición de x no es única. Supongamos que es única y determinemos la intersección de E_1 y E_2; sea x un elemento común a E_1 y E_2, se tendrá

$$x = x + 0 \qquad x \in E_1 \qquad 0 \in E_2$$
$$x = 0 + x \qquad 0 \in E_1 \qquad x \in E_2$$

por tanto, $x = 0$, pues la descomposición de todo elemento x debe ser único, luego

$$E_1 \cap E_2 = \{0\}.$$

Recíprocamente, supongamos $E = E_1 + E_2$, $E_1 \cap E_2 = \{0\}$, o sea,

$$x = x_1 + x_2 = x_1' + x_2' \qquad x_1, x_1' \in E_1 \qquad x_2, x_2' \in E_2$$
$$x_1 - x_1' = x_2' - x_2 = 0$$

pues este elemento $x_1 - x_1' = x_2' - x_2$ pertenece a $E_1 \cap E_2$, de donde:

Teorema 2 y definición. — *Dados dos subespacios vectoriales* E_1, E_2 *de un espacio vectorial* E *tales que* $E = E_1 + E_2$, *las dos condiciones siguientes son equivalentes:*

1. *La descomposición de todo elemento x de E en suma* $x_1 + x_2$, x_1 *perteneciente a* E_1 *y* x_2 *a* E_2 *es único.*

2. $E_1 \cap E_2 = \{0\}$.

Cuando una de estas condiciones se cumple, se dice que E_1 *y* E_2 *son* dos subespacios suplementarios *de* E *y se escribe*

$$E = E_1 \oplus E_2.$$

Entonces surgen dos preguntas: si E_1 es un subespacio de E, *¿existe un* suplementario E_2 de E_1 respecto a E, y si existe es *único*? La respuesta a la segunda pregunta es negativa (ver ej. 1 más abajo). Respecto a la primera la respuesta es positiva cualquiera que sea el subespacio E_1 de E; existe un suplementario, como lo demostraremos en el § 137 (teorema 9) para los espacios de dimensión finita.

OBSERVACION

No se confundirá *suplementario* del subespacio E_1 del espacio vectorial E y el *complementario* del conjunto E_1 respecto a E: por otra parte, $\complement_E E_1$ no es un subespacio de E, pues no contiene el 0.

c) Si E_1 y E_2 son dos subespacios suplementarios de E, consideremos la descomposición única de todo x de E

$$x = x_1 + x_2, \qquad x_1 \in E_1, \qquad x_2 \in E_2$$

y la aplicación de $E_1 \times E_2$ en $E = E_1 \oplus E_2$ definida por

$$(x_1, x_2) \to f(x_1, x_2) = x_1 + x_2$$

es *biyectiva*, pues para todo x de E, si la descomposición $x = x_1 + x_2$ es única, $f(x_1, x_2) = x$ tiene una solución única (x_1, x_2); por otra parte, cualquiera que sean (x_1, x_2) e (y_1, y_2) de $E_1 \times E_2$ y α de K

$$f[(x_1, x_2) + (y_1, y_2)] = f(x_1 + y_1, x_2 + y_2) = (x_1 + y_1) + (x_2 + y_2)$$
$$= (x_1 + x_2) + (y_1 + y_2) = f(x_1, x_2) + f(y_1, y_2)$$
$$f[\alpha(x_1, x_2)] = f(\alpha x_1, \alpha x_2) = \alpha x_1 + \alpha x_2 = \alpha(x_1 + x_2) = \alpha f(x_1, x_2)$$

por tanto:

TEOREMA 3.—*Si* E_1 *y* E_2 *son dos subespacios vectoriales suplementarios de* E, *el espacio vectorial* $E_1 \times E_2$ *es isomorfo a* $E = E_1 \oplus E_2$.

Recíprocamente, dados dos espacios vectoriales E_1 y E_2 sobre K para todo $x = (x_1, x_2)$ de $E = E_1 \times E_2$, se tiene de una manera única

$$(x_1, x_2) = (x_1, 0) + (0, x_2);$$

por tanto, $E = E_1 \times E_2 = E'_1 \oplus E'_2$, E'_1 (resp. E'_2) subespacio de E descrito por $(x_1, 0)$ (resp. $(0, x_2)$) es isomorfo a E_1 (resp. a E_2) (ver § 128, ej. 5).

Identificando E'_1 y E_1 (resp. E'_2 y E_2) se ve que E_1 y E_2 son *subespacios suplementarios* de $E_1 \times E_2$. Esto equivale en definitiva a identificar los dos espacios isomorfos $E_1 \times E_2$ y $E_1 \oplus E_2$.

En estas condiciones, si se tiene para todo x de $E = E_1 \oplus E_2$, la descomposición única $x = x_1 + x_2$, la aplicación de E en E_1 definida por

$$x = x_1 + x_2 \to x_1$$

está identificada con la aplicación de $E = E_1 \times E_2$ en E_1 que es la proyección pr_1 de $E_1 \times E_2$ sobre E_1, se le llama la *proyección de* E *sobre el subespacio* E_1, *paralelamente al subespacio* E_2 *suplementario de* E_1 *respecto a* E. *Se dice que* x_1 (resp. x_2) *es la componente de* x *de* $E_1 \oplus E_2$ *en* E_1 (resp. E_2).

d) Sea E_1 un subespacio vectorial de E, E_2 el suplemento de E_1 respecto a E. La relación de equivalencia $x - y \in E_1$ se escribe con las notaciones evidentes

$$(x_1 + x_2) - (y_1 + y_2) = x_1 - y_1 + x_2 - y_2 \in E_1$$

es decir, $x_2 = y_2$. Por tanto, todo representante $\dot{x}$ de una clase módulo E_1, de E/E_1 tiene la misma proyección x_2 sobre E_2 paralelamente a E_1. Consideremos la aplicación f de E/E_1 en E_2 definida por

$$\dot{x} = \dot{\widehat{x_1 + x_2}} \to f(\dot{x}) = x_2$$

es *suprayectiva*, pues cualquiera que sea x_2 de E_2, existe $\dot{x}_2$, tal que $f(\dot{x}_2) = x_2$; por otra parte, $f(\dot{x}) = f(\dot{x}') = x_2$ implica $x = x_1 + x_2$, $x' = x'_1 + x_2$; por tanto, $x - x' \in E_1$, es decir, $\dot{x} = \dot{x}'$, f es *inyectiva*. En fin, si $y = y_1 + y_2$

$$f(\dot{x} + \dot{y}) = f\left(\dot{\widehat{x+y}}\right) = x_2 + y_2 = f(\dot{x}) + f(\dot{y})$$

y para todo α de K

$$f(\alpha\dot{x}) = f\left(\dot{\widehat{\alpha x}}\right) = \alpha x_2 = \alpha f(\dot{x})$$

por tanto:

TEOREMA 4. — *Si* E_1 *y* E_2 *son subespacios suplementarios del espacio vectorial* E, *el espacio cociente* E/E_1 *es isomorfo a* E_2.

COROLARIO. — *En un espacio vectorial* E, *todos los suplementarios de un subespacio vectorial* E_1 *son isomorfos.*

EJEMPLOS Y EJERCICIOS

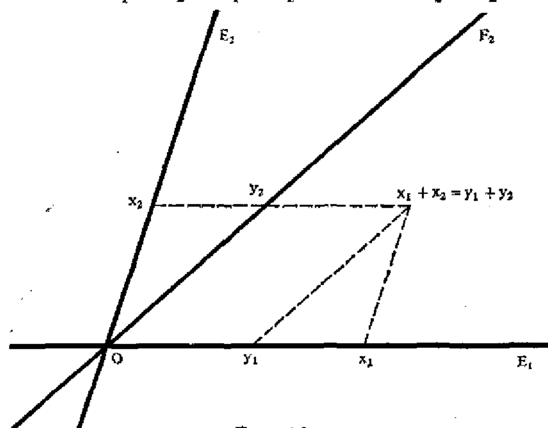
1. Sea E_1 el espacio vectorial de los vectores del eje real Ox_1 y E_2 el espacio vectorial de los vectores libres del eje real Ox_2, distinto de Ox_1; consideremos $E_1 \times E_2$ isomorfo a $\mathbf{R}_2$. Identificar E_1 y E'_1 quiere decir confundir x_1 vector de E_1 y $(x_1, 0)$ vector de $E_1 \times E_2$ e identificar $E_1 \times E_2$ y $E_1 \oplus E_2$ equivale a confundir (x_1, x_2) de $E_1 \times E_2$ y $x_1 + x_2$ de $E_1 \oplus E_2$. x_1 es la proyección de $x = x_1 + x_2$ sobre E_1 paralelamente a E_2: tal es el origen de la definición general dada al final del subpárrafo *c*).

Consideremos un tercer eje real, Oy_2 distinto de Ox_1 y Ox_2, sea F_2 el espacio vectorial de los vectores libres de Oy_2; tendremos de una manera única para todo x de E

$$x = x_1 + x_2, \quad x_1 \in E_1, \quad x_2 \in E_2$$
$$x = y_1 + y_2, \quad y_1 \in E_1, \quad y_2 \in F_2.$$

Luego

$$E = E_1 \oplus E_2 = E_1 \oplus F_2 \quad \text{con} \quad E_2 \neq F_2.$$



2. En $E = \mathfrak{F}(\mathbf{R}, \mathbf{R})$ (§ 125, ej. 5), sea P el conjunto de las funciones pares e I el conjunto de las funciones impares, demostrar que $E = P \oplus I$. Se demostrará que para toda función f de E y todo t real se tiene

$$f(t) = \frac{1}{2}[f(t) + f(-t)] + \frac{1}{2}[f(t) - f(-t)].$$

3. Sea $\mathfrak{P}$ el espacio vectorial sobre **R** de los polinomios con coeficientes reales (§ 125, ej. 2, y capítulo 11) y A un polinomio de grado superior o igual a 1. Si E_1 es el conjunto de los polinomios múltiplos del polinomio A y E_2 el conjunto de los polinomios de grado estrictamente inferior al de A, demostrar que E_1 y E_2 son subespacios suplementarios de $\mathfrak{P}$. (Utilizar la división euclídea de los polinomios.)

4. Si E_1, E_2 son dos subespacios de E, ¿en qué caso $E_1 \cup E_2$ es un subespacio de E?

132. Suma y suma directa de varios subespacios vectoriales

a) Dados n subespacios vectoriales $E_1, E_2, \ldots, E_n$ de un espacio vectorial E, se verificará fácilmente que el conjunto representado $E_1 + E_2 + \ldots + E_n$ descrito por $x_1 + x_2 + \ldots + x_n$, donde x_i describe E_i $(1 \leq i \leq n)$ es un subespacio vectorial de E; es el menor subespacio que contiene cada E_i, por tanto:

TEOREMA Y DEFINICIÓN. — *Si* (E_i) $(1 \leq i \leq n)$ *es una familia finita de subespacios vectoriales de* E, *el conjunto* $E_1 + E_2 + \ldots + E_n$ *es el subespacio engendrado por* $E_1 \cup E_2 \cup \ldots \cup E_n$; *se le llama* subespacio suma *de los subespacios* $E_1, E_2, \ldots, E_n$.

b) Se puede generalizar la noción de subespacios suplementarios:

DEFINICIÓN. — *Se dice que el espacio vectorial* E *es la* suma directa *de los subespacios* $E_1, E_2, \ldots, E_n$, *si todo* x *de* E *puede escribirse de un modo único en la forma*

$$x = x_1 + x_2 + \ldots + x_n,$$

donde x_i *pertenece a* E_i $(1 \leq i \leq n)$. *Se escribe*

$$E = E_1 \oplus E_2 \oplus \ldots \oplus E_n.$$

Se dice que x_i *es la* componente *de* x *en* E_i.

Si E_1 *y* E_2 *son dos subespacios suplementarios de* E, *se dirá también que* E *es suma directa de* E_1 *y* E_2.

EJERCICIOS

1. Demostrar que E es suma directa de $E_1, E_2, \ldots, E_n$ si y sólo si

$$E = E_1 + E_2 + \ldots + E_n$$

y, además,

$$\forall i \in [1, n-1] \qquad (E_1 + E_2 + \ldots + E_i) \cap E_{i+1} = \{0\}.$$

(Se observará que esta segunda condición es más fuerte que la condición

$$i \neq j \Rightarrow E_i \cap E_j = \{0\}.)$$

2. Si $E = E_1 \oplus E_2 \oplus \ldots \oplus E_n$, demostrar que E es isomorfo a $E_1 \times E_2 \times \ldots \times E_n$.

III. Independencia lineal. Bases

133. Independencia lineal

a) Sea E un espacio vectorial sobre K de dimensión finita y $G = \{g_1, g_2, \ldots, g_n\}$ una *parte generatriz* de E (§ 129, *b*). Es evidente que toda parte G′ de E que contiene G es también una parte generatriz de E; inversamente se puede preguntar si existen partes propias G″ de G que engendran también E; para simplificar, se puede igualmente hallar una *parte generatriz minimal de* E, *es decir, una parte generatriz* G *de* E *tal que para cada uno de sus elementos* g_i, $G - \{g_i\}$ *no engendre* E. Supongamos que suceda así. Observemos que un tal sistema minimal es forzosamente tal que $g_1, g_2, \ldots, g_n$ sean todos distintos y no nulos; vamos también a ver que una relación de la forma[28]

$$\lambda^1 g_1 + \lambda^2 g_2 + \ldots + \lambda^n g_n = 0$$

sólo es posible si $\lambda^1 = \lambda^2 = \ldots = \lambda^n = 0$; supongamos que una de las λ^i sea no nula, cambiando si es necesario la numeración $i \to g_i$, se puede siempre suponer que es el coeficiente λ^n de g_n; λ^n es entonces *inversible* en K y se obtiene

$$g_n = -\lambda^1(\lambda^n)^{-1}g_1 + \ldots - \lambda^{n-1}(\lambda^n)g_{n-1} = \mu^1 g_1 + \ldots + \mu^{n-1}g_{n-1}$$

pero entonces para todo *a* de E, tendremos

$$\begin{aligned} a &= \alpha^1 g_1 + \alpha^2 g_2 + \ldots + \alpha^{n-1}g_{n-1} + \alpha^n g_n \\ &= \alpha^1 g_1 + \alpha^2 g_2 + \ldots + \alpha^{n-1}g_{n-1} + \alpha^n(\mu^1 g_1 + \ldots + \mu^{n-1}g_{n-1}) \end{aligned}$$

y $\{g_1, g_2, \ldots, g_{n-1}\} = G - \{g_n\}$ sería una parte generatriz de E, lo que es contrario a la hipótesis.

Estas consideraciones nos conducen a dar la definición siguiente en un espacio vectorial cualquiera:

Definición. — *Una familia finita* (x_i) $(i \in [1, p])$ *de elementos de un espacio vectorial* E *es* libre *si*

$$\lambda^1 x_1 + \lambda^2 x_2 + \ldots + \lambda^p x_p = 0 \Rightarrow (\lambda^1 = \lambda^2 = \ldots = \lambda^p = 0).$$

Una familia cualquiera (x_i) $(i \in I)$ *es* libre *si todas sus subfamilias finitas son libres. Una familia que no es libre se le llama* ligada.

Por consiguiente, una familia cualquiera es ligada si hay una subfamilia finita ligada y una familia finita (x_i) $(i \in [1, p])$ es ligada si existen $\lambda^1, \lambda^2, \ldots, \lambda^p$, elementos de K, *no todos nulos* tales que

$$\lambda^1 x_1 + \lambda^2 x_2 + \ldots + \lambda^p x_p = 0.$$

---

(28) Recuérdese que en λ^i y después en μ^i, α^i, ..., *i* es un *índice* y no un exponente. Por el contrario en $(\lambda^n)^{-1}$, -1 es un exponente; si se tiene que emplear un exponente, lo que será raro, salvo -1, se escribirá $(\lambda^i)^k$.

De lo anterior resulta que los elementos de una *familia libre* son todos distintos: en efecto, si se tenía $h \neq k$ y $x_h = x_k$, la subfamilia $(h, k) \rightarrow (x_h, x_k)$ sería ligada. Por consiguiente, si la familia (x_i) $(i \in \mathrm{I})$ es libre, la aplicación f de I sobre $f(\mathrm{I})$ definida por $i \rightarrow f(i) = x_i$ es *biyectiva*: se puede confundir sin inconveniente (ver § 17, observación) la parte $f(\mathrm{I})$ de E y la familia (x_i) $(i \in \mathrm{I})$, diremos que los elementos de una familia libre describen una *parte libre* de E, o también que los elementos de una familia libre son *linealmente independientes*. Una parte no libre de E se llama *parte ligada* de E, se dice que sus elementos son *linealmente dependientes*.

b) De las definiciones resultan las *propiedades* inmediatas siguientes:

— *Toda subfamilia de una familia libre es libre.*

— *Toda superfamilia de una familia ligada es ligada.*

— *Los elementos de una familia libre son no nulos,* pues si, por ejemplo, $x_p = 0$, se tendría con $\lambda \neq 0$

$$0x_1 + 0x_2 + \ldots + 0x_{p-1} + \lambda x_p = 0.$$

En particular, $\{x\}$ es libre si y sólo si $x \neq 0$.

— *Si una parte* $\{x_1, x_2, \ldots, x_p\}$ *que pertenece a un subespacio vectorial* F *de* E, *es libre* (resp. *ligada*) *en* F, *es libre* (resp. *ligada*) *en* E.

— *Sea* $\mathrm{E} = \mathrm{E}' \oplus \mathrm{E}''$, $x_i = x_i' + x_i''$, $x_i' \in \mathrm{E}'$ y $x_i'' \in \mathrm{E}''$.

Si $\{x_1, x_2, \ldots, x_p\}$ *es ligada en* E, $\{x_1', x_2', \ldots, x_p'\}$ *es ligada en* E'; *si* $\{x_1', x_2', \ldots, x_p'\}$ *es libre en* E', $\{x_1, x_2, \ldots, x_p\}$ *es libre en* E.

Atención: *Si* $\{x_1, x_2, \ldots, x_p\}$ *es libre en* E, *no se puede decir nada de la parte* $\{x_1', x_2', \ldots, x_p'\}$ *proyectada sobre* E' *paralelamente a* E''.

c) Teorema 1. — *Una familia es ligada si y sólo si existe un elemento en la familia que sea combinación lineal finita de los otros elementos de la familia.*

Según una observación hecha anteriormente, sólo hay que tener en cuenta una familia finita.

La proposición directa es evidente. Recíprocamente sea una familia (x_i) $(1 \leq i \leq p)$ ligada, existe una familia de escalares $\lambda^1, \lambda^2, \ldots, \lambda^p$ *no todos nulos* tales que

$$\lambda^1 x_1 + \lambda^2 x_2 + \ldots + \lambda^p x_p = 0$$

supongamos, cambiando si es preciso la numeración, que $\lambda^p \neq 0$, luego $(\lambda^p)^{-1}$ existe, de donde

$$\begin{aligned} x_p &= -\lambda^1(\lambda^p)^{-1}x_1 - \lambda^2(\lambda^p)^{-1}x_2 - \ldots - \lambda^{p-1}(\lambda^p)^{-1}x_{p-1} \\ &= \mu^1 x_1 + \mu^2 x_2 + \ldots + \mu^{p-1}x_{p-1}. \end{aligned}$$

En particular, si una familia con dos elementos no nulos es ligada, existe un escalar λ tal que $x_2 = \lambda x_1$: se dice que x_1 y x_2 son *colineales*.

COROLARIO. — *Si* $\{x_1, x_2, \ldots, x_p\}$ *es una parte libre de* p *elementos de* E *y* $\{x_1, x_2, \ldots, x_p, x\}$ *una parte ligada,* x *pertenece al subespacio engendrado por* $\{x_1, x_2, \ldots, x_p\}$ *y se tiene*

$$x = \mu^1 x_1 + \ldots + \mu^p x_p$$

de una manera única.

Tenemos, en efecto,

$$\lambda^1 x_1 + \lambda^2 x_2 + \ldots + \lambda^p x_p + \lambda x = 0$$

con $\lambda^1, \lambda^2, \ldots, \lambda^p, \lambda$ no todos nulos. $\lambda \neq 0$, pues si se tuviera $\lambda = 0$ uno de los escalares $\lambda^1, \lambda^2, \ldots, \lambda^p$ sería no nulo; por tanto, la parte $\{x_1, x_2, \ldots, x_p\}$ sería ligada. Como en la demostración del teorema 1, se tendrá

$$x = \mu^1 x_1 + \ldots + \mu^p x_p.$$

Supongamos

$$\mu^1 x_1 + \ldots + \mu^p x_p = \mu'^1 x_1 + \ldots + \mu'^p x_p,$$

se tendrá

$$(\mu^1 - \mu'^1) x_1 + \ldots + (\mu^p - \mu'^p) x_p = 0$$

por tanto, $\mu^i - \mu'^i = 0$ $(1 \leq i \leq p)$, si la parte $\{x_1, \ldots, x_p\}$ es libre.

d) TEOREMA 2. — *Si* $L = \{a_1, a_2, \ldots, a_m\}$ *es una parte libre con* m *elementos de un espacio vectorial* E *y* $G = \{g_1, g_2, \ldots, g_p\}$ *es una parte generatriz de* E *con* p *elementos,* $m \leq p$ *(y si se cambia eventualmente de numeración* $i \to g_i$*),* $G' = \{a_1, a_2, \ldots, a_m, g_{m+1}, \ldots, g_p\}$ *es también una parte generatriz de* E.

En efecto(29),

$$a_1 = \alpha_1^1 g_1 + \ldots + \alpha_p^1 g_p$$

al menos un α_1^i es no nulo si no $a_1 = 0$ y L no sería libre, cambiando eventualmente de numeración $i \to g_i$, podemos suponer $\alpha_1^1 \neq 0$; por tanto,

$$g_1 = \beta_1^1 a_1 + \beta_1^2 g_2 + \ldots + \beta_1^p g_p.$$

De lo anterior resulta que $G_1 = \{a_1, g_2, g_3, \ldots, g_p\}$ es una parte generatriz de E; por tanto,

$$a_2 = \alpha_2^1 a_1 + \alpha_2^2 g_2 + \ldots + \alpha_2^p g_p$$

al menos uno de los escalares $\alpha_2^2 \ldots \alpha_2^p$ es no nulo, pues si $\alpha_2^2 = \ldots = \alpha_2^p = 0$, se tendría $a_2 = \alpha_2^1 a_1$ y L no sería libre; sea, cambiando si es preciso de numeración. $\alpha_2^2 \neq 0$, tendremos

$$g_2 = \beta_2^1 a_1 + \beta_2^2 a_2 + \beta_2^3 g_3 + \ldots + \beta_2^p g_p$$

---

(29) Cada a_i de E es una combinación lineal de $g_1, g_2, \ldots, g_p$; en esta combinación lineal el coeficiente de g_j, sea α_i^j, tendrá dos índices i inferior, número de a_i, y j índice superior, número de g_j; por tanto, α_i^j es el coeficiente. Se tiene $a_i = \sum_{j=1}^{p} \alpha_i^j g_j$, en el segundo miembro j puede reemplazarse por cualquier índice (salvo i, 1, p); por otra parte, como no figura en el primer miembro, se dice que es un *índice mudo* (ver § 6, *b*, nota 2).

por tanto, $G_2 = \{a_1, a_2, g_3, \ldots, g_p\}$ es también una parte generatriz de E; si se repite un número finito de veces esta operación, vemos que para $p' \leq \inf(m, p)$

$$G_{p'} = \{a_1, a_2, \ldots, a_{p'}, g_{1+p'}, \ldots, g_p\}$$

es una parte generatriz de E.

Demostremos que $m > p$ es imposible, en este caso haciendo $p' = p$

$$G_p = \{a_1, a_2, \ldots, a_p\}$$

engendraría E y los elementos $a_{p+1}, a_{p+2}, \ldots, a_m$ serían combinaciones lineales de $a_1, a_2, \ldots, a_p$; por tanto, L no sería libre.

Por consiguiente, $m \leq p$ y al cabo de un número finito de operaciones, llegaremos, por tanto, a la parte generatriz

$$G' = G_m = \{a_1, a_2, \ldots, a_m, g_{m+1}, \ldots, g_p\}.$$

De $m \leq p$ resulta el corolario siguiente:

Corolario. — *Si* G *es una parte generatriz, con* p *elementos, de un espacio vectorial* E, *toda parte de* E *que tiene estrictamente más de* p *elementos es ligada.*

Lo que se puede expresar en la forma equivalente:

Toda parte de $p + 1$ *vectores de* E *que son combinación lineal de* p *vectores cualesquiera de* E *es ligada.*

EJEMPLOS Y EJERCICIOS

1. Consideremos un vector $x_i = (\alpha_i^1, \ldots, \alpha_i^j, \ldots, \alpha_i^n)$ de $\mathbf{K}^n$; en el escalar α_i^j, i y j son dos índices, i índice inferior es el índice del vector x_i, j índice superior es el índice que indica el número de la coordenada α_i^j.

Sea el vector de $\mathbf{K}^n$, $a_i = \{\delta_i^1, \delta_i^2, \ldots, \delta_i^j, \ldots, \delta_i^n\}$ tal que

$$\delta_i^j = 0 \quad \text{si} \quad i \neq j$$
$$\delta_i^j = 1 \quad \text{si} \quad i = j$$

δ_i^j se llama *símbolo de Kronecker*. Los elementos $a_1, a_2, \ldots, a_n$ son linealmente independientes; en efecto,

$$\lambda^1 a_1 + \lambda^2 a_2 + \ldots + \lambda^n a_n = 0_{\mathbf{K}^n} = (0, 0, \ldots, 0)$$

implica para todo j de $[1, n]$

$$\lambda^1 \delta_1^j + \ldots + \lambda^i \delta_i^j + \ldots + \lambda^n \delta_n^j = 0$$

es decir,

$$\lambda^j = 0.$$

2. En el espacio vectorial de los polinomios en x con coeficientes reales (o complejos), toda familia finita o infinita extraída de $\{x^0, x^1, x^2, \ldots, x^n, \ldots\}$ es libre (ver capítulo 11).

3. En el espacio vectorial de las aplicaciones de **R** en **R** las aplicaciones $t \to e^{rt}$, $t \to e^{st}$ son linealmente independientes si y sólo si los reales r y s son distintos (ver generalización, ej. 148 al final del capítulo).

4. Si E es un espacio vectorial sobre **K** designado E_K, se considera E como espacio vectorial $E_{K'}$ sobre $K' \subset K$. Estudiar la independencia lineal de una familia $x_1, x_2, \ldots, x_p$ según que se le considere que pertenece a E_K o a $E_{K'}$. Dar ejemplos considerando el cuerpo **C** ya como espacio vectorial sobre sí mismo (§ 125, ej. 3), ya como espacio vectorial sobre **R** (§ 125, ej. 4).

134. Bases de un espacio vectorial de dimensión finita

Sea un espacio vectorial E sobre el cuerpo K, de dimensión finita, y $G_m = \{g_1, g_2, \ldots, g_p\}$ una *parte generatriz minimal de* E que tiene p elementos (ver § 133, *a*), hemos visto que esta parte es libre; es, por otra parte, la noción de parte generatriz minimal la que nos ha servido de introducción a la noción de independencia lineal.

Recíprocamente, sea $B = \{a_1, a_2, \ldots, a_p\}$ una *parte generatriz libre* cualquiera, con p elementos, vamos a demostrar que es una parte generatriz minimal; en efecto, si no lo fuera, uno de estos elementos sería una combinación lineal de los otros y B no sería una parte libre (ver § 133, *c*).

Por otro lado, si x es un elemento cualquiera de E y $B = \{a_1, a_2, \ldots, a_p\}$ una parte generatriz libre (por tanto, minimal), $\{a_1, a_2, \ldots, a_p, x\}$ es una parte ligada, sino $\{a_1, a_2, \ldots, a_p\}$ no engendraría E; B es, por tanto, una familia libre tal que si se le añade un elemento cualquiera x, cesa de serlo: se dice que es una *parte libre maximal.*

Recíprocamente, sea $L_M = \{b_1, b_2, \ldots, b_n\}$ una parte libre maximal, es decir, una parte libre tal que para todo x de E, $L_M \cup \{x\}$ sea ligada; según el corolario del teorema 1 del § 133, todo x de E es una combinación lineal de $b_1, b_2, \ldots, b_n$; por tanto, $\{b_1, b_2, \ldots, b_n\}$ es una parte generatriz, por otro lado, libre, de E, de donde:

TEOREMA 3 Y DEFINICIÓN. — *Para una parte* B *no vacía de un espacio vectorial sobre* K, *de dimensión finita, las tres propiedades siguientes son equivalentes:*

1. B *es una parte generatriz libre de* E.
2. B *es una parte generatriz minimal de* E.
3. B *es una parte libre maximal de* E.

Toda parte $B = \{a_1, a_2, \ldots, a_n\}$ *que posee una de estas propiedades se le llama una* base *de* E. *Para todo* x *de* E *existe una familia única de escalares* (α^i) $(1 \leq i \leq n)$ *tales que*

$$x = \alpha^1 a_1 + \alpha^2 a_2 + \ldots + \alpha^n a_n = \sum_{i=1}^{n} \alpha^i a_i$$

$\alpha^1, \alpha^2, \ldots, \alpha^n$ *se llaman las* **coordenadas** *de x respecto a la base* $\{a_1, a_2, \ldots, a_n\}$.

Observemos que el hecho de que B no sea vacío implica que $E \neq \{0\}$.

OBSERVACION

En lugar de coordenadas de x algunos autores emplean el término componentes, es mejor reservar el término de componente para el vector $\alpha^i a_i$; en efecto, si E_i es el subespacio engendrado por a_i, $E = E_1 \oplus E_2 \oplus \dots \oplus E_n$ y $\alpha^i a_i$ es la componente de x en E_i (ver § 132, *b*).

EJEMPLOS

1. En K^n hemos visto (§ 133, ej. 1) que para $1 \leq i \leq n$ el sistema de las a_i definidas por $a_i = (\delta_i^1, \dots, \delta_i^i, \dots, \delta_i^n)$ es un sistema libre; por otro lado, para todo x de K^n

$$x = (\alpha^1, \alpha^2, \dots, \alpha^i, \dots, \alpha^n) = \sum_{i=1}^{n} (0, 0, \dots, 0, \alpha^i, 0, \dots, 0) = \sum_{i=1}^{n} \alpha^i a_i;$$

por tanto, la familia (a_i) $(1 \leq i \leq n)$ es una familia generatriz libre de K^n; es, por tanto, una *base* de K^n, y se le llama la *base canónica* de K^n, espacio vectorial sobre K.

2. El teorema 3 es también cierto para los espacios vectoriales que no son de dimensión finita. Por ejemplo, en el espacio vectorial $\mathcal{P}$ de los polinomios en x con coeficientes reales (o complejos), $\{x^0, x^1, \dots, x^n, \dots\}$ es una base de $\mathcal{P}$ (aquí 0, 1, ..., n, ... son exponentes) (ver capítulo 11).

Todo elemento A de $\mathcal{P}$ es una combinación lineal finita (§ 128, *b*) de elementos de la base, pero el número de elementos de la base que intervienen con los coeficientes no nulos en la expresión de un polinomio A no está acotado cuando A describe $\mathcal{P}$.

135. Existencia de bases para un espacio de dimensión finita

a) Acabamos de dar tres caracterizaciones equivalentes de una base de un espacio vectorial de dimensión finita sobre un cuerpo K; vamos ahora a demostrar su existencia.

Sea E un espacio vectorial de dimensión finita sobre K no reducido a $\{0\}$, admite por definición una parte generatriz finita $G = \{g_1, g_2, \dots, g_p\}$ de p elementos que se puede suponer no nulos según la observación hecha en el § 129, *b*). Hay, pues, partes de G que son libres (por ejemplo, $\{g_1\}$) sea L una de ellas; por tanto, $L \subset G$.

Si L engendra E, es una base (parte generatriz libre). Si L no engendra E, existe g_{i_1} de $G - L$ que no pertenece al subespacio F_L engendrado por L, pues si todos los elementos de $G - L$ pertenecieran a F_L, L engendraría E; pongamos

$$L_0 = L, \qquad L_1 = L_0 \cup \{g_{i_1}\}$$

L_1 es libre, si no (corolario del teorema 1, § 133) g_{i_1} pertenecería a F_L, luego

$$L = L_0 \subset L_1 \subset G \qquad L_1 \neq L_0.$$

Si L_1 engendra E, L_1 es una base de E. Si L_1 no engendra E, podemos empezar de nuevo el razonamiento: existe g_{i_2} que pertenece a $G - L_1$, $L_2 = L_1 \cup \{g_{i_2}\}$ es libre y

$$L_0 \subset L_1 \subset L_2 \subset G$$

con las dos primeras inclusiones estrictas, podemos construir así una *sucesión finita estrictamente creciente* de partes libres de E, contenidas todas en G

$$L = L_0 \subset L_1 \subset L_2 \subset \dots \subset L_k \subset G.$$

Al ser G una parte generatriz finita, existirá m tal que L_{m-1} sea una parte libre que no engendra E y L_m una parte libre que engendra E, L_m será, por tanto, una base de E; luego:

TEOREMA 4. — *Todo espacio vectorial* E *de dimensión finita, no reducido a* $\{0\}$, *admite una base; de una manera más precisa si* G *es una parte generatriz de* E *y* L *una parte libre de* E *contenida en* G, *existe una base* B *de* E *tal que*

$$L \subset B \subset G.$$

b) Consideremos ahora un espacio vectorial E no reducido a $\{0\}$, una parte libre L y una parte generatriz G, las dos cualesquiera. $G \cup L$ engendra *a fortiori* E y $L \subset G \cup L$, existe (teorema 4) una base B tal que

$$L \subset B \subset G \cup L.$$

Todos estos conjuntos al ser finitos, existe una parte H de G tal que $B = L \cup H$; en consecuencia:

COROLARIO. — *Si* L *y* G *son, respectivamente, una parte libre y una parte generatriz de un espacio vectorial* E, *existe una parte* H *de* G *tal que* $L \cup H$ *sea una base de* E.

Este resultado se conoce con el nombre de "*teorema de la base incompleta*": la parte libre L se ha "*completado*" con algunos elementos de G; también se le conoce con el nombre de "*teorema de cambio*", pues gracias al teorema 2 (§ 133) se puede sustituir a $G = \{g_1, g_2, \dots, g_p\}$ por la parte generatriz $G' = \{a_1, a_2, \dots, a_m, g_{m+1}, \dots, g_p\}$ conteniendo L obtenida "cambiando" m elementos de G por los m elementos de la parte libre L.

OBSERVACION

Hemos demostrado los teoremas 3 y 4 solamente para los espacios vectoriales de dimensión finita; también se verifican para un espacio vectorial cualquiera.

136. Dimensión de un espacio vectorial

a) Todo espacio vectorial E de dimensión finita admite al menos una base finita B (§ 135, teorema 4). Sea n el número de sus elementos. Sea B′ otra base que tenga n' elementos. B′ es una parte libre de E y sus elementos son combinaciones lineales de los elementos de B; por tanto (§ 133, corolario del teorema 2), $n' \leq n$, igualmente $n \leq n'$. Luego:

TEOREMA 5 Y DEFINICIÓN. — *En un espacio vectorial de dimensión finita sobre el cuerpo* K, *todas las bases tienen el mismo número de elementos. Este número común se llama la* dimensión *del espacio vectorial* E *sobre el cuerpo* K, *se le designa* $\dim_K E$.

Por ejemplo, K^n espacio vectorial sobre K, tiene una base con n elementos; por ejemplo, su base canónica es de dimensión n sobre K (§ 134, ej. 1).

Así está justificada *a posteriori* la calificación de *espacio vectorial de dimensión finita*. Se demuestra que el teorema 5 se verifica también para los espacios vectoriales cualesquiera en el sentido de que todas las bases de un espacio vectorial son equipotentes.

Como para la independencia lineal (ver § 133, ej. 4), la noción de *dimensión* depende no solamente del conjunto E, sino también del *cuerpo de base* del espacio vectorial E (ver ej. 2 a continuación). Si no es posible ninguna confusión sobre el cuerpo de base, se representará la dimensión de E por dim E.

$\{0\}$ es un espacio vectorial con un solo elemento cualquiera que sea el cuerpo de base, se pondrá por definición

$$\dim \{0\} = 0.$$

b) TEOREMA 6. — *Todo espacio vectorial* E *de dimensión n sobre* K *es isomorfo a* K^n *(espacio vectorial sobre* K*).*

Sea $\{a_1, a_2, \ldots, a_n\}$ una base de E, para todo x de E

$$x = \alpha^1 a_1 + \ldots + \alpha^n a_n.$$

Consideremos la aplicación E en K^n definida por

$$x \to (\alpha^1, \alpha^2, \ldots, \alpha^n)$$

es claramente una biyección; por otra parte, si

$$y \to (\beta^1, \beta^2, \ldots, \beta^n)$$
$$x + y \to (\alpha^1 + \beta^1, \alpha^2 + \beta^2, \ldots, \alpha^n + \beta^n) = (\alpha^1, \alpha^2, \ldots, \alpha^n) + (\beta^1, \beta^2, \ldots, \beta^n)$$
$$\lambda x \to (\lambda\alpha^1, \lambda\alpha^2, \ldots, \lambda\alpha^n) = \lambda(\alpha^1, \alpha^2, \ldots, \alpha^n)$$

según las propiedades del espacio producto K^n (§ 127). Por tanto, esta aplicación es un *isomorfismo* de E sobre K^n. Es decir, existe una sola estructura de espacio vectorial de dimensión n sobre K: la de K^n.

COROLARIO. — *Dos espacios vectoriales de dimensión finita sobre el mismo cuerpo* K *son isomorfos si y sólo si tienen la misma dimensión respecto a* K.

Observemos que *este isomorfismo de* E *sobre* K^n *no es canónico,* pues depende no sólo del espacio vectorial E, sino también de la base elegida.

c) De la definición de la dimensión y del corolario del teorema 2 (§ 133) tenemos los resultados siguientes:

TEOREMA 7. — *En un espacio vectorial* E *de dimensión n sobre* K:

1. *Toda parte libre tiene a lo sumo n elementos.*
2. *Toda parte que tenga al menos* $n + 1$ *elementos es ligada.*

COROLARIO. — *Toda parte* B *de un espacio vectorial* E *de dimensión n sobre* K *que posee dos de las tres propiedades siguientes* ($n > 0$):

1. B *tiene n elementos.*
2. B *es libre.*
3. B *engendra* E, *es una base de* E.

2. y 3. dan la definición de una base.

3. y 1. implican que B es libre, si no se podría extraer B′ parte estricta de B (§ 135, teorema 4) que sería una base de E y se tendría dim E $< n$.

1. y 2. implican que B engendra E, si no existiría x tal que B $\cup \{x\}$ sería libre, y habría entonces una parte libre con $n + 1$ elementos, lo que es imposible.

EJEMPLOS Y EJERCICIOS

1. Todo cuerpo K es un espacio de dimensión 1 sobre sí mismo, si a es un elemento no nulo cualquiera de K, para todo x de K existe α tal que $x = \alpha a$ ($\alpha = xa^{-1}$); por tanto, $\{a\}$ es una base.

2. Consideremos el conjunto **C** de los números complejos, podemos considerarlo como un espacio vectorial sobre **R**, tendrá por base $1 = (1, 0)$ e $i = (0, 1)$, o todo par de números complejos linealmente independientes en **C** espacio vectorial sobre **R**, es decir, tal que $z_1/z_2 \notin \mathbf{R}$; entonces tendremos de un modo único para todo z de **C**

$$z = \alpha_1 z_1 + \alpha_2 z_2.$$

Se puede también considerar el cuerpo **C** como espacio vectorial sobre sí mismo; tiene por base toda parte de un elemento $\{1\}$, por ejemplo, o $\{z_0\}$ si $z_0 \neq 0$ y para todo z de **C** $z = \alpha z_0$.

Por tanto,

$$\dim_{\mathbf{R}} \mathbf{C} = 2 \qquad \dim_{\mathbf{C}} \mathbf{C} = 1.$$

3. Un espacio vectorial tal como el espacio vectorial $\mathcal{P}$ de los polinomios en x con coeficientes reales (o complejos) es de dimensión infinita. El conjunto $\mathcal{P}_n$ de los polinomios en x con coeficientes reales (o complejos) de grado estrictamente inferior a n admite por base

$$\{x^0, x^1, x^2, \ldots, x^{n-1}\}$$

es, pues, de dimensión n (ver capítulo 11).

4. Si E y F son dos espacios vectoriales sobre K de dimensiones respectivas finitas p y q, demostrar que si (a_i) $(1 \leq i \leq p)$ es una base de E y (b_j) $(1 \leq j \leq q)$ una base de F, $((a_i, 0)) \cup ((0, b_j))$, donde i describe $[1, p]$ y j, $[1, q]$ es una base de E $\times$ F, deducir de lo anterior que dim (E $\times$ F) = dim E + dim F.

Generalizar a n espacios vectoriales, de dimensiones finitas sobre K

$$\dim (E_1 \times E_2 \times \ldots \times E_n) = \dim E_1 + \dim E_2 + \ldots + \dim E_n.$$

5. Sea E de dimensión finita $n > 0$ sobre K. Demostrar que E tiene una infinidad de elementos si y sólo si K es infinito. ¿Cuál es el cardinal de E si card K = p?

137. Dimensión de un subespacio vectorial de E

a) Sea E un espacio vectorial de dimensión finita $n > 0$ sobre K (por tanto, E $\neq \{0\}$) y F un subespacio de E, distinto de $\{0\}$. Recordemos que toda parte libre L de F es una parte libre de E (§ 133, b) y tiene, por tanto, a lo sumo n elementos (§ 136, teorema 7); L tiene al menos un elemento no nulo, puesto que F $\neq \{0\}$; hay, pues, en F partes libres, sea p el número de elementos de *una parte libre maximal* de F: es una base de F

(§ 134) y $1 \leq p \leq n$; si $p = n$ es una base de E y E = F. Si F = {0}, $\dim_K F = 0$. Finalmente si E = {0}, se tiene igualmente F = {0}, de donde:

TEOREMA 8. — *Si* F *es un subespacio vectorial del espacio vectorial* E *de dimensión* n *sobre* K, F *es de dimensión finita sobre* K *y*

$$\dim_K F \leq \dim_K E.$$

Recíprocamente toda parte libre con p elementos de E *engendra un subespacio vectorial de* E, *de dimensión p. Finalmente*

$$\dim_K F = \dim_K E \Rightarrow F = E.$$

b) Un subespacio F de dimensión 1 se llama una *recta* que pasa por 0 de E, si a es un elemento no nulo de F, F está descrito por las x tales que

$$x = \alpha a$$

cuando α describe K.

Un subespacio F de dimensión 2 se llama un *plano* que pasa por 0 de E, si $\{a_1, a_2\}$ es una base de F, F está descrito por los x tales que

$$x = \alpha^1 a_1 + \alpha^2 a_2$$

si (α^1, α^2) describe K × K.

Un subespacio F de dimensión $p > 2$ de base $\{a_1, a_2, \ldots, a_p\}$ está descrito por los x tales que

$$x = \sum_{i=1}^{p} \alpha^i a_i$$

si $(\alpha^1, \alpha^2, \ldots, \alpha^p)$ describe K^p. Si $p = n - 1$ se dice que F es un ***hiperplano*** que pasa por 0 de E.

OBSERVACION

Naturalmente todos los subespacios vectoriales de E contienen 0; se les dice recta, plano, hiperplano, que *pasan por* 0 para evitar toda confusión con recta, plano, hiperplano del espacio afín asociado al espacio vectorial E (ver tomo III Geometría(*)). Por ejemplo, en el espacio afín $\mathbf{R}^3$ hay rectas y planos que no pasan por 0.

c) Consideremos un subespacio propio F del espacio vectorial E, sea $\{a_1, a_2, \ldots, a_p\}$ $(0 < p < n)$ una base de F, podemos completar esta base con elementos de una parte generatriz de E para obtener una base de E (§ 135, corolario del teorema 4); como una base de E tiene n elementos será necesario $n - p$ elementos $b_{p+1}, b_{p+2}, \ldots, b_n$. Para todo x de E existirá, pues, una familia única de escalares

$$\alpha^1, \alpha^2, \ldots, \alpha^p, \beta^{p+1}, \ldots, \beta^n$$

tales que

$$x = \alpha^1 a_1 + \alpha^2 a_2 + \ldots + \alpha^p a_p + \beta^{p+1} b_{p+1} + \ldots + \beta^n b_n.$$

(*) *N. del T.* — El tomo III anunciado no ha sido publicado aún en la Colección U de la A. Colin.

Ahora bien, $\{b_{p+1}, b_{p+2}, \ldots, b_n\}$ subfamilia de una familia libre de E es libre y engendra un subespacio vectorial G de dimensión $n-p$, pues su base tiene $n-p$ elementos. Por otra parte, escribiendo

$$y = \alpha^1 a_1 + \ldots + \alpha^p a_p, \qquad z = \beta^{p+1} b_{p+1} + \ldots + \beta^n b_n$$

vemos que tenemos de un modo único

$$x = y + z, \qquad y \in \mathrm{F}, \qquad z \in \mathrm{G}$$

por tanto, G es un subespacio suplementario de F, de donde:

TEOREMA 9. — *En un espacio vectorial* E *de dimensión finita sobre* K *todo subespacio vectorial* F *de* E *admite al menos un suplementario* G *respecto a* E *y*

$$\mathrm{E} = \mathrm{F} \oplus \mathrm{G} \Rightarrow \dim \mathrm{E} = \dim \mathrm{F} + \dim \mathrm{G}.$$

Este resultado es, en efecto, válido para $\mathrm{F} = \{0\}$ (resp. E), G es entonces E (resp. $\{0\}$).

Observemos que F puede tener varios suplementarios (§ 131, ej. 1), puede incluso tener infinitos si K es infinito (ver ej. 4 más abajo).

COROLARIO Y DEFINICIÓN. — *En un espacio vectorial* E *de dimensión finita sobre* K *todos los suplementarios de un mismo subespacio vectorial* F *de* E *tienen igual dimensión, se le llama la* codimensión *de* F *respecto a* E *y se la designa* $\mathrm{codim}_{\mathrm{E}}\, \mathrm{F}$.

Volvemos a encontrar así, gracias a la noción de dimensión, que todos los suplementarios de un mismo espacio vectorial en un espacio vectorial de dimensión finita son isomorfos. Lo habíamos demostrado en el § 131 (corolario del teorema 4) para un espacio vectorial cualquiera al admitir que un subespacio vectorial tiene suplementarios.

EJERCICIOS

1. Demostrar que si E es de dimensión finita sobre K

$$\mathrm{E} = \mathrm{E}_1 \oplus \mathrm{E}_2 \oplus \ldots \oplus \mathrm{E}_n \Rightarrow \dim \mathrm{E} = \sum_{i=1}^{n} \dim \mathrm{E}_i.$$

2. Si tenemos dos subespacios vectoriales de E de dimensión finita, demostrar que

$$\dim (\mathrm{E}_1 + \mathrm{E}_2) + \dim (\mathrm{E}_1 \cap \mathrm{E}_2) = \dim \mathrm{E}_1 + \dim \mathrm{E}_2.$$

(Introducir $\mathrm{I} = \mathrm{E}_1 \cap \mathrm{E}_2$, un suplementario F_1 de I respecto a E_1 y un suplementario de F_2 de I respecto a E_2. Considerando las bases de I, F_1 y F_2, demostrar que $\mathrm{E}_1 + \mathrm{E}_2 = \mathrm{I} \oplus \mathrm{F}_1 \oplus \mathrm{F}_2$.)

Se indicarán otras demostraciones en el § 143 (ej. 1) y en el ejercicio 164, al final del capítulo.

3. En E de dimensión n sobre K, se considera dos subespacios vectoriales de dimensiones, respectivamente, p_1 y q_2. ¿Entre qué límites pueden variar $\dim (\mathrm{E}_1 + \mathrm{E}_2)$ y $\dim (\mathrm{E}_1 \cap \mathrm{E}_2)$?

4. Sea E un espacio de dimensión finita, $\mathrm{E} = \mathrm{F} \oplus \mathrm{G}$, $\mathrm{A} = \{a_1, a_2, \ldots, a_p\}$ una base de F y $\mathrm{B} = \{b_1, b_2, \ldots, b_q\}$ una base de G. Si a' es un elemento no nulo de F, se pone

$B' = \{b'_1, \ldots, b'_q\}$ con $b'_j = a' + b_j$, demostrar que $A \cup B'$ es una base de E y que B' engendra un subespacio G' suplementario de F y distinto de G.

Demostrar que si se sustituye a' por a'' elemento no nulo de F y distinto de a', se obtiene un suplementario G'' de E distinto de G'

Deducir que si K es infinito, F tiene una infinidad de suplementarios (utilizar § 136, ej. 5).

138. Rango de un sistema de vectores de un espacio vectorial

DEFINICIÓN. — *Se llama rango de un sistema* S *de vectores de un espacio vectorial* E *sobre* K, *la dimensión del subespacio* F, *supuesto de dimensión finita, engendrado por este sistema de vectores. Se le representa* rg (S).

Dado $S = \{x_1, x_2, \ldots, x_p\}$ que engendra F, el rango de S será el número de elementos de una parte maximal libre extraída de S, por tanto:

El rango de un sistema finito de vectores es el número máximo de vectores linealmente independientes extraídos de S.

Si dim $E = n$, se tiene evidentemente

$$\text{rg}\,\{x_1, x_2, \ldots, x_p\} \leq \inf\,(n, p).$$

Vamos a indicar un primer método para determinar el rango de S, en un espacio de E de dimensión n sobre K, si conocemos las coordenadas de los p vectores respecto a una base de E, sea

$$x_i = \alpha_i^1 a_1 + \alpha_i^2 a_2 + \ldots + \alpha_i^j a_j + \ldots + \alpha_i^n a_n$$

será lo mismo que estudiar los vectores $(\alpha_i^1, \alpha_i^2, \ldots, \alpha_i^j, \ldots, \alpha_i^n)$ de K^n. Esto se basa en dos propiedades muy simples:

PROPIEDAD 1. — *Si tenemos* p *vectores* $x_i = \sum_{j=1}^{n} \alpha_i^j a_j$ *de un espacio vectorial* E *de dimensión* n *sobre* K $(p \leq n)$, *si*

$$(j < i \Rightarrow \alpha_i^j = 0) \qquad \text{y} \qquad (\forall\, i \in [1, p]) \qquad \alpha_i^i \neq 0$$

los p *vectores* (x_i) *son independientes.*

Sea

$$\lambda^1 x_1 + \ldots + \lambda^i x_i + \ldots + \lambda^p x_p = 0$$

considerando las coordenadas de $\sum_{i=1}^{p} \lambda^i x_i$ tenemos

$$\sum_{i=1}^{p} \lambda^i \alpha_i^j = 0 \qquad (j = 1, 2, \ldots, p)$$

o sea, si simplificamos y tenemos en cuenta las propiedades de las α_i^j

$$\begin{aligned}
&\lambda^1\alpha_1^1 = 0\\
&\lambda^1\alpha_2^2 + \lambda^2\alpha_2^2 = 0\\
&\cdot\\
&\cdot\\
&\cdot\\
&\lambda^1\alpha_1^j + \lambda^2\alpha_2^j + \dots + \lambda^j\alpha_j^j = 0\\
&\cdot\\
&\cdot\\
&\cdot\\
&\lambda^1\alpha_1^p + \lambda^2\alpha_2^p + \dots + \lambda^p\alpha_p^p = 0
\end{aligned}$$

lo que nos da

$$\lambda^1 = \lambda^2 = \dots = \lambda^p = 0.$$

PROPIEDAD 2. — *El sistema* $S = \{x_1, x_2, \dots, x_p\}$ *y el sistema* S' *obtenido sustituyendo en* S, x_i *por* $\sum_{j=1}^{p} \lambda^j x_j$ *con* $\lambda^i \neq 0$ *tienen el mismo rango.*

En efecto, el subespacio engendrado por S y por

$$S' = (S - \{x_i\}) \cup \{\lambda^1 x_1 + \dots + \lambda^i x_i + \dots + \lambda^p x_p\}$$

es el mismo si $\lambda^i \neq 0$.

El método consiste en utilizar la propiedad (2) efectuando combinaciones sucesivas $\lambda^i x_i + \lambda^j x_j$ para obtener un sistema que engendre el mismo subespacio que S y cuyos vectores no nulos posean la propiedad (1). Vamos a exponerlo con ejemplos:

EJEMPLOS

1. Sean los tres vectores de $\mathbf{R}^3$ (o de un espacio vectorial isomorfo), x_1, x_2, x_3 dados por sus coordenadas en una base $\{a_1, a_2, a_3\}$

(S)

x_1	x_2	x_3
2	—1	4
3	2	—3
5	—3	—8.

Escojamos uno de los vectores cuya primera coordenada es no nula (si las tres primeras coordenadas fueran nulas se operaría en el subespacio engendrado por a_2 y a_3, al que pertenecerían x_1, x_2, x_3). Tomemos x_1 y formemos $S' = \{x_1, x'_2, x'_3\}$ con $x'_i = \lambda x_1 + \mu x_i$ $(i = 2, 3)$, λ y μ se escogen, cada vez, de manera que la primera coordenada de x'_2 y x'_3 sea nula, obtenemos

(S′)

x_1	$x'_2 = x_1 + 2x_3$	$x'_3 = 2x_1 - x_3$
2	0	0
3	7	9
5	—1	2

Formemos seguidamente $S'' = \{x_1, x_2', x_3''\}$ con $x_3'' = \lambda x_2' + \mu x_3'$, se han escogido λ y μ de manera que la segunda coordenada de x_3'' sea nula, obtenemos

(S'')

x_1	x_2'	$x_3'' = 9x_2' - 7x_3'$
2	0	0
3	7	0
5	—1	—23

Los tres sistemas S, S', S'' tienen el mismo rango (propiedad 2), S'' está formado por tres vectores independientes (propiedad 1); por tanto, el rango de S es 3.

2. En el mismo espacio vectorial que en el ejemplo 1, consideremos el sistema $T = S \cup \{x_4\}$, donde x_4 tiene por coordenadas —4, 17, —10, obtenemos los sistemas T, T', T'' que tienen el mismo rango

(T')

x_1	x_2'	x_3'	$x_4' = 2x_1 + x_4$
2	0	0	0
3	7	9	23
5	—1	2	0

(T'')

x_1	x_2'	x_3''	$x_4'' = 23x_2' - 7x_3'$
2	0	0	0
3	7	0	0
5	—1	—23	—23

Se podría formar el sistema $T''' = \{x_1, x_2', x_3'', x_4'''\}$ con $x_4''' = x_3'' - x_4'' = 0$: el sistema T''', por ser el sistema T de rango 3, tiene los cuatro vectores dependientes, lo que no es extraño, puesto que la dimensión del espacio vectorial es 3. El método nos da la relación de dependencia —aquí única— verificada por x_1, x_2, x_3, x_4. En efecto,

$$0 = x_3'' - x_4'' = 9x_2' - 7x_3' - (23x_2' - 7x_4') = -7(2x_2' + x_3' - x_4')$$

nos da

$$x_4 = 2x_1 + 4x_2 - x_3.$$

En todos los casos comprobaremos que el método nos da no solamente el rango r del subespacio F engendrado por $S = \{x_1, x_2, \ldots, x_p\}$, sino también una base de F y las $p - r$ relaciones de dependencia existentes entre $x_1, x_2, \ldots, x_p$. En particular, si en E de dimensión n, $x_1, x_2, \ldots, x_n$ son independientes, este método permite calcular las coordenadas de un vector x cualquiera respecto a la base $\{x_1, x_2, \ldots, x_n\}$.

IV. Propiedades de las aplicaciones lineales

Dada la importancia de las aplicaciones lineales, agrupamos en esta sección todos los resultados referentes a ellas sin temor a repetir definiciones o demostraciones ya dadas, o dadas en parte.

139. Definiciones. Ejemplos

DEFINICIÓN. — *Dados dos espacios vectoriales* E *y* F *sobre el mismo cuerpo conmutativo* K, *se llama* aplicación lineal *de* E *en* F *todo* homomorfismo *de* E *en* F, *es decir, una aplicación* f *de* E *en* F, *tal que*

(1) $(\forall x \in E)$ $(\forall y \in E)$ $f(x + y) = f(x) + f(y)$

(2) $(\forall x \in E)$ $(\forall \alpha \in K)$ $f(\alpha x) = \alpha f(x)$.

(1) y (2) se llaman, respectivamente, *primera* y *segunda* propiedades de *linealidad.* Si E = F se dice que f es un *endomorfismo* del espacio vectorial E, se dice también que es un *operador lineal* operando en E.

Si la aplicación lineal f: E $\to$ F es *biyectiva,* es un *isomorfismo* de E sobre F; si, además, E = F, f es un *automorfismo* del espacio vectorial E, se dice también que f es un *operador lineal regular* operando en E.

Se representa por $\mathrm{Hom}_K(E, F)$ o $\mathcal{L}_K(E, F)$ el conjunto de las aplicaciones lineales de E en F, espacios vectoriales sobre K, $\mathrm{End}_K(E)$ o $\mathcal{L}_K(E)$ el conjunto de los endomorfismos del espacio vectorial E y $\mathbf{GL}_K(E)$ el conjunto de los automorfismos de E; si no cabe ninguna confusión sobre el cuerpo de base K, se emplean las notaciones Hom(E, F), End(E) y **GL**(E); de hecho estas notaciones se aplican a los conjuntos de aplicaciones lineales provistos de *operaciones algebraicas* que definiremos en la sección V (§§ 144, 146 y 147).

EJEMPLOS Y EJERCICIOS

1. Si $E = E_1 \times E_2$ las aplicaciones pr_1 y pr_2 definidas por

$$pr_1(x_1, x_2) = x_1, \qquad pr_2(x_1, x_2) = x_2$$

son aplicaciones lineales de $E_1 \times E_2$, respectivamente, en E_1 y en E_2 (si E_1 y E_2 son espacios vectoriales sobre K).

Si se identifica $E = E_1 \times E_2$ y $E_1 \oplus E_2$ (§ 131, *c*) podemos decir que pr_1 y pr_2 son endomorfismos de E. Considerando $E_1 \times E_2 \times \ldots \times E_n$ o $E_1 \oplus E_2 \oplus \ldots \oplus E_n = E$ ($x = x_1 + x_2 + \ldots + x_n$, $x \in E$, $x_i \in E_i$) se constatará que $x \to x_i = f_i(x)$ es un endomorfismo de E. Se comprueba fácilmente que $(f_i) \circ (f_i) = f_i$, de una manera general se llama *proyector* todo endomorfismo de E que verifica $f \circ f = f$ (ver ej. 157, al final del capítulo).

2. Si F es un subespacio vectorial del espacio vectorial E sobre K, la aplicación f de E en E/F definido por

$$x \to f(x) = \dot{x}$$

($\dot{x}$ clase módulo F de elementos de E) es un homomorfismo *suprayectivo,* llamado *homomorfismo canónico* de E sobre E/F (§ 130).

3. Sea **E** el conjunto de las funciones numéricas indefinidamente derivables $t \to x(t)$ de [0, 1] en **R**, es un espacio vectorial sobre **R**.

a) La aplicación de E en E que a la función de x hace corresponder su derivada x', es un endomorfismo de E (ver curso de Análisis).

b) La aplicación de E en **R** definida por $x \to x'(t_0)$ (t_0 valor fijo de [0, 1] es una aplicación lineal de E en **R**.

4. Sea E el conjunto de las funciones numéricas continuas $t \to x(t)$ de [0, 1] en **R**, es un espacio vectorial sobre **R**.

a) La aplicación de E en E que a la función x hace corresponder su primitiva X que se anula para $t = 0$, o sea, $t \to X(t) = \int_0^t x(u)\,dx$ es un endomorfismo de E (ver curso de Análisis).

b) La aplicación de E en **R** definida por $x \to \int_0^1 x(t)\,dt$ es una aplicación lineal.

5. Se volverá a ver todos los isomorfismos vistos en los párrafos precedentes (§ 125, ej. 4 y 8; § 128, ej. 5; § 131, t. 3 y 4; § 132, ej. 2; § 136, t. 6).

140. Propiedades fundamentales de las aplicaciones lineales

Hay que tener siempre en cuenta que una aplicación lineal f de un espacio vectorial E en un espacio vectorial F (sobre el mismo cuerpo K) es un *homomorfismo del grupo aditivo* E *en el grupo aditivo* F.

Volvemos a encontrar así en este párrafo tres teoremas (t. 1, 2, 3) análogos a los ya demostrados para los homomorfismos de grupos (§ 77) y como corolario la descomposición canónica de una aplicación lineal. Sin embargo, volvemos a hacer todas las demostraciones.

Las nociones de subespacios suplementarios y de independencia lineal nos darán en los §§ 141 y 142 dos nuevos teoremas (t. 4 y 5). Aunque sea válido para los espacios vectoriales de dimensión infinita, el teorema 4 que supone la existencia de suplementarios para un subespacio vectorial de E, sólo se demostrará en este curso más que para los espacios vectoriales de dimensión finita sobre K.

Finalmente, en el § 143, nos limitaremos a los espacios de dimensión finita sobre K para los teoremas 6 y 7 (determinación y rango de una aplicación lineal).

a) TEOREMA 1. — *La aplicación compuesta de dos aplicaciones lineales es una aplicación lineal.*

Si E, F, G son tres espacios vectoriales sobre el mismo cuerpo K, consideremos dos aplicaciones lineales f de E en F y g de F en G; $g \circ f$ es una aplicación de E en G. Para todo x y todo y de E

$$\begin{aligned} (g \circ f)(x+y) &= g[f(x+y)] && \text{(definición de } g\circ f) \\ &= g[f(x)+f(y)] && (f \in \mathcal{L}(\mathrm{E},\ \mathrm{F})) \\ &= g[f(x)] + g[f(y)] && (g \in \mathcal{L}(\mathrm{E},\ \mathrm{F})) \\ &= (g\circ f)(x) + (g \circ f)(y) && \text{(definición de } g\circ f). \end{aligned}$$

Igualmente y por las mismas razones, para todo x de E y todo α de K

$$(g \circ f)(\alpha x) = g[f(\alpha x)] = g[\alpha f(x)] = \alpha g[f(x)] = \alpha(g \circ f)(x).$$

COROLARIO. — *La composición de dos isomorfismos de espacios vectoriales es un isomorfismo de espacios vectoriales. La composición de dos endomorfismos* (resp. *automorfismos*) *de un espacio vectorial* E *es un endomorfismo* (resp. *automorfismo*) *de* E.

TEOREMA 2. — *Si* f *es una aplicación de* E *en* F

1. $f(0_E) = 0_F$, $f(-x) = -f(x)$.
2. *Si* A *es un subespacio vectorial de* E, f(A) *es un subespacio vectorial de* F.
3. *Si* B *es un subespacio vectorial de* F, f^{-1}(B) *es un subespacio vectorial de* E.

Recordemos la demostración de 1 (§ 56, t. 2). Sea x' de f(E), existe x de E tal que $f(x) = x$

$$x + 0_E = 0_E + x = x \Rightarrow f(x + 0_E) = f(x) + f(0_E) = f(0_E) + f(x) = f(x)$$

o

$$x' + f(0_E) = f(0_E) + x' = x',$$

es decir,

$$f(0_E) = 0_F.$$

Finalmente

$$x + (-x) = 0_E \Rightarrow f(x) + f(-x) = f(0_E) = 0_F$$

por tanto, $f(-x) = -f(x)$.

Sea A un subespacio de E, x' e y' dos elementos *cualesquiera* de $f(A)$, existe x e y de A tales que $x' = f(x)$, $y' = f(y)$, de donde para todo α de K

$$\begin{cases} f(x - y) = f(x) - f(y) = x' - y' \\ f(\alpha x) = \alpha f(x) = \alpha x' \end{cases}$$

$x - y$ y αx si pertenecen a A, $x' - y'$ y $\alpha x'$ pertenecerán a $f(A)$, que es, por tanto, un subespacio de F (§ 128).

Sea B un subespacio de F, x e y dos elementos *cualesquiera* de $f^{-1}(B)$, $f(x)$ y $f(y)$ pertenecen a B, de donde para todo α de K

$$\begin{cases} f(x) - f(y) = f(x - y) \\ \alpha f(x) = f(\alpha x) \end{cases}$$

pertenecen a B, luego $x - y$ y αx pertenecen a $f^{-1}(B)$ que es, por tanto, un subespacio de E.

DEFINICIÓN. — *Si f es una aplicación lineal de* E *en* F, $f^{-1}(0)$, *subespacio vectorial de* E *se llama* núcleo *de la aplicación lineal f y se le representa* Ker f; $f(E)$ *subespacio vectorial de* F *se le llama* imagen *de la aplicación lineal f y se representa* Im f.

$$f \in \mathcal{L}(E, F), \qquad \text{Ker } f = f^{-1}(0) \subset E, \qquad \text{Im } f = f(E) \subset F.$$

COROLARIO. — *Si f es una aplicación lineal de* E *en* F:

1. *f es inyectiva si y sólo si* Ker $f = \{0\}$.
2. *f es suprayectiva si y sólo si* Im f = F.

Para 1: $f(x) = f(y) \Rightarrow f(x - y) = 0$, es decir, $x - y \in \text{Ker } f$, luego $x = y$ si y sólo si Ker $f = \{0\}$.

2 es simplemente el enunciado de la definición de una aplicación suprayectiva.

TEOREMA 3. — *Si f es un isomorfismo del espacio vectorial* E *sobre el espacio vectorial* F, *f^{-1} es un isomorfismo de* F *sobre* E.

Para todo x' e y' de F existe x e y únicos de E tal que $f(x) = x'$, $f(y) = y'$, de donde para todo α de K

$$f(x + y) = f(x) + f(y) = x' + y' \Rightarrow f^{-1}(x' + y') = x + y = f^{-1}(x') + f^{-1}(y')$$
$$f(\alpha x) = \alpha f(x) = \alpha x' \qquad \Rightarrow f^{-1}(\alpha x) = \alpha x = \alpha f^{-1}(x').$$

b) Como para los grupos si consideramos la relación de equivalencia (§ 19, *b*) definida sobre E por $f(x) = f(y)$, tendremos

$$f(x - y) = 0 \Leftrightarrow x - y \in f^{-1}(0)$$

pues si $f^{-1}(0) = N$ es un subespacio vectorial de E, esta relación es *compatible con la estructura del espacio vectorial* de E y la descomposición

$$E \xrightarrow{s} E/N \xrightarrow{b} f(E) \xrightarrow{i} E$$
$$f = i \circ b \circ s$$

es tal que E/N y $f(E)$ son los espacios vectoriales sobre K; s es suprayectiva. Por otra parte, para todo x e y de E y todo α de K (ver § 130, y § 139, ej. 2)

$$\begin{cases} s(x+y) = \left(\dot{\widehat{x+y}}\right) = \dot{x} + \dot{y} = s(x) + s(y) \\ s(\alpha x) = \left(\dot{\widehat{\alpha x}}\right) = \alpha(\dot{x})\, \alpha s(x) \end{cases}$$

$\dot{x}$ e $\dot{y}$ son las clases módulo N, respectivamente, de x e y; s es, por tanto, un *homomorfismo suprayectivo*. Por otra parte, b es una biyección definida por (§ 19, b) para todo $\dot{x}$ de E/N

$$b(\dot{x}) = f(x).$$

Para todo $\dot{x}$ y todo $\dot{y}$ de E/N y para todo α de K

$$\begin{cases} b(\dot{x} + \dot{y}) = b\left[\dot{\widehat{x+y}}\right] = f(x+y) = f(x) + f(y) = b(\dot{x}) + b(\dot{y}) \\ b(\alpha\dot{x}) = b\left[\dot{\widehat{\alpha x}}\right] = f(\alpha x) = \alpha f(x) = \alpha b(\dot{x}) \end{cases}$$

luego b es un *isomorfismo* de E/N sobre $f(E)$. Finalmente la aplicación canónica de $f(E)$ en F, $x' \to x'$ es visiblemente lineal, es *inyectiva*; por tanto:

Corolario. — *Toda aplicación lineal f: $E \to F$, si E y F son dos espacios vectoriales sobre K, se descompone de una manera canónica en* (N *es el núcleo de* f):

a) *El homomorfismo canónico de* E *sobre* E/N.

b) *El isomorfismo canónico de* E/N *sobre* $f(E)$.

c) *El homomorfismo canónico (inyectivo) de* $f(E)$ *en* F.

EJERCICIOS

1. ¿Cuál es el núcleo de f: $E_1 \times E_2 \to E_1$ (§ 139, ej. 1)? Hallar así de nuevo la causa de que $E_1 \oplus E_2/E_1$ sea isomorfo a E_2 (§ 131, t. 4).

2. ¿Cuál es el núcleo de las aplicaciones lineales definidas en los ejemplos 3 *a*) y 4 *a*) del § 139?

3. Si f es una aplicación lineal de E en F, demostrar que la restricción g de f a un subespacio vectorial E′ de E es una aplicación lineal de E′ en F y que $\operatorname{Ker} g = \operatorname{Ker} f \cap E'$.

141. Isomorfismo de Im f con todo suplementario de Ker f

TEOREMA 4. — *Sea f una aplicación lineal de un espacio vectorial* E *sobre un espacio vectorial* F, *si* E_2 *es un suplementario de* $E_1 = \text{Ker } f$ *respecto a* E, $f(E_2) = f(E)$ *y la restricción* g *de* f *a* E_2 *que toma sus valores en* $f(E)$ *es un isomorfismo de* E_2 *sobre* $f(E)$.

Como ya hemos dicho, este teorema válido para un espacio vectorial E cualquiera sobre K sólo se halla demostrado en este curso para los espacios vectoriales de dimensión finita sobre K.

Para todo x de $E = E_1 \oplus E_2$ ($E_1 = \text{Ker } f$) tenemos la descomposición única

$$x = x_1 + x_2, \qquad x_1 \in \text{Ker } f \Rightarrow f(x_1) = 0, \qquad x_2 \in E_2$$

de donde para todo x de E

$$f(x) = f(x_1 + x_2) = f(x_1) + f(x_2) = f(x_2)$$

por tanto, $f(E_2) = f(E)$. Sea g la restricción f a E_2 y que toma sus valores en $f(E)$. g es *suprayectiva*, pues para todo x' de $f(E)$ existe $x = x_1 + x_2$ de E ($x_1 \in E_1$ $x_2 \in E_2$) tal que

$$x' = f(x) = f(x_1 + x_2) = f(x_2) = g(x_2)$$

por otra parte, g es inyectiva, pues

$$g(x_2) = g(y_2) \Rightarrow g(x_2 - y_2) = 0$$

por tanto, $x_2 - y_2$ que pertenece por definición a E_2 pertenece también al núcleo de f, puesto que $g(x_2 - y_2) = f(x_2 - y_2) = 0$, luego $x_2 - y_2 = 0$, puesto que E_1 y E_2 son suplementarios; g es, por tanto, una *biyección* de E_2 sobre $f(E)$.

Finalmente g es una *aplicación lineal,* pues para todo x_2, todo y_2 de E y todo α de K

$$\begin{cases} g(x_2 + y_2) = f(x_2 + y_2) = f(x_2) + f(y_2) = g(x_2) + g(y_2) \\ g(\alpha x_2) = f(\alpha x_2) = \alpha f(x_2) = \alpha g(x_2). \end{cases}$$

OBSERVACIONES

1. Hemos demostrado en un caso aparentemente particular el ejercicio 3 del § 140.

Por otra parte, el hecho que todo suplementario del núcleo es isomorfo a $f(E)$ es una consecuencia del teorema 4 del § 131 ($E_1 \oplus E_2/E_1$ *isomorfo a* E_2) y del corolario de los teoremas 1, 2 y 3 de este párrafo (E/N isomorfo a $f(E)$); además, la demostración dada no hace sino reconsiderar lo esencial de las demostraciones de estos resultados.

2. Si $E_1 = \text{Ker } f$ y $E = E_1 \oplus E_2$ acabamos de demostrar que

$$f(E_1) = \{0\}, \qquad f(E_2) = f(E); \qquad \text{por tanto,} \qquad f(E) = \{0\} \oplus f(E).$$

Hay que tener en cuenta que el resultado no es general, es decir, si E es suma directa de dos subespacios vectoriales cualesquiera E_1 y E_2, $f(E)$ no es en general suma directa de $f(E_1)$ y $f(E_2)$ (ver ej. fin del § 142).

142. Imágenes de partes de E. Aplicaciones

Sea f una aplicación lineal de E en F y una combinación lineal de p elementos de E: $x_1, \ldots, x_p$, tendremos

$$f\left(\sum_{i=1}^{p} \alpha^i x_i\right) = \sum_{i=1}^{p} f(\alpha^i x_i) \qquad \text{(primera propiedad de linealidad)}$$

$$= \sum_{i=1}^{p} \alpha^i f(x_i) \qquad \text{(segunda propiedad de linealidad).}$$

Igualmente para toda combinación lineal finita de una familia cualquiera de E

$$f\left(\sum_{i\in I} \alpha^i x_i\right) = \sum_{i\in I} f(\alpha^i x_i) = \sum_{i\in I} \alpha^i f(x_i)$$

recordemos que en este último caso sólo un número finito de escalares α^i son no nulos. Si A es una parte de E, bien finita $\{x_1, x_2, \ldots, x_p\}$ o infinita (x_i) con I como conjunto de índices, representaremos como siempre por $f(A)$ el conjunto de las imágenes $f(x_i)$ de las x_i por f.

Observemos que si la aplicación $i \to x$ es biyectiva, no lo es en general la $i \to f(x_i)$; este fenómeno explica, en parte, algunos de los resultados siguientes:

TEOREMA 5.— *Si f es una aplicación lineal de* E *en* F, *si* G *es una parte generatriz de* E, *$f(G)$ es una parte generatriz de $f(E)$.*

Sea x' un elemento cualquiera de $f(E)$, existe x de E tal que $x' = f(x)$. Por otra parte, x es una combinación lineal finita de elementos de G y

$$x = \sum_{i\in I} \alpha^i x_i \Rightarrow x' = f(x) = \sum_{i\in I} \alpha^i f(x_i),$$

por tanto, $f(x)$ es una combinación lineal finita de elementos de $f(G)$.

En particular, si $f(x) = 0$ para todo x de G, $f(x) = 0$ para todo x de E: esta observación, muy útil, se conoce con el nombre de *principio de prolongación de igualdades lineales*. Por ejemplo, si f y g son dos aplicaciones lineales de E en F, tales que $f(x) = g(x)$ para todo x de G, se deduce que $f = g$.

COROLARIO 1.— *Si f es una aplicación lineal de* E *en* F, *la imagen de una familia* A *ligada de* E *es una familia $f(A)$ ligada de* F.

Según el teorema 1 (§ 133) existe un elemento x de A combinación lineal finita de los elementos de $A - \{x\}$; por tanto, $f(x)$ será combinación lineal finita de los elementos de $f(A) - \{f(x)\}$, luego $f(A)$ es ligada.

COROLARIO 2. — *Si f es una aplicación lineal de* E *en* F *y* A *una familia de* E *tal que* $f(A)$ *sea libre en* F, *entonces* A *es libre en* E.

En efecto, si A no fuera libre en E, $f(A)$ tampoco sería libre en F.

ATENCIÓN: La imagen de una familia libre de E no es en general una familia libre de F (ver ej. posterior); en particular, la imagen de una base de E no es en general una base de $f(E)$.

EJERCICIO

Si f es una aplicación lineal de E en F, demostrar que las tres propiedades siguientes son equivalentes:

a) f es inyectiva.

b) Para toda familia libre L de E, $f(L)$ es una familia libre de F.

c) Para toda descomposición $E = E_1 \oplus E_2$ se tiene $f(E) = f(E_1) \oplus f(E_2)$.

143. Caso en que E es de dimensión finita. Rango de una aplicación lineal de E de dimensión finita en F

a) Si E y F son dos espacios vectoriales sobre K, supongamos $\dim_K E = n$ y designamos por $\{a_1, a_2, \ldots, a_n\}$ una base de E.

TEOREMA 6. — *Existe una aplicación lineal única f de un espacio vectorial* E *de dimensión n sobre* K *en un espacio vectorial* F *sobre* K, *tal que*

$$(\forall i \in [1, n]) \qquad f(a_i) = b_i$$

donde $b_1, b_2, \ldots, b_n$ son n elementos cualesquiera de F.

Supongamos que existe f y verifica las condiciones del enunciado, tendremos de una manera única

$$x = \sum_{i=1}^{n} \alpha^i a_i \Rightarrow f(x) = \sum_{i=1}^{n} \alpha^i f(a_i) = \sum_{i=1}^{n} \alpha^i b_i;$$

por tanto, si f existe, es única. Recíprocamente, la aplicación de E en F definida por

$$x \to \sum_{i=1}^{n} \alpha^i b_i$$

es evidentemente lineal. Por tanto, f existe, y es única, está *determinada por sus valores sobre los elementos de la base.*

b) TEOREMA 7 Y DEFINICIÓN. — *Si f es una aplicación lineal de un espacio vectorial* E *de dimensión finita sobre* K *en un espacio vectorial* F *sobre* K, $f(E)$ *es de dimensión finita sobre* K *y* $\dim_K f(E) \leq \dim_K E$.

La dimensión de Im $f = f(E)$ *es el* rango *de la aplicación lineal f; se representa por* rg (f); *además,*

$$\boxed{\operatorname{rg}(f) = \dim_K E - \dim_K \operatorname{Ker} f.}$$

Si $\dim_K E = n$, toda base B de E tiene n elementos, es una parte generatriz de E; por tanto, $f(B)$ es una parte generatriz de $f(E)$ que tiene a lo sumo n elementos. Por tanto, $f(E)$ es de dimensión finita y toda base de $f(E)$, parte generatriz minimal de $f(E)$ (§ 134, teorema 3), tiene $r \leq n$ elementos.

Por otra parte, si $E_1 = \text{Ker } f$ y $E = E_1 \oplus E_2$, tendremos (§ 137)

$$\dim_K E = \dim_K \text{Ker } f + \dim E_2;$$

ahora bien, E_2 es isomorfo a $f(E)$; por tanto, $\dim_K E_2 = \dim_K f(E)$ y

$$r = \text{rg } (f) = \dim_K E - \dim_K \text{Ker } f.$$

Se observará que este resultado es *independiente del hecho de que* F *sea de dimensión infinita o finita y en este último caso del valor de su dimensión.* Sin embargo, F es de dimensión finita p

$$\text{rg } (f) \leq \inf (n, p),$$

pues $f(E) \subset F$.

COROLARIO 1. — *Sea f una aplicación lineal de* E *en* F, E *y* F *dos espacios vectoriales de dimensiones finitas sobre* K, *respectivamente, iguales a n y p:*

1. $\text{rg } (f) = n$ *si y sólo si f es inyectiva.*
2. $\text{rg } (f) = p$ *si y sólo si f es suprayectiva.*

En efecto, $\text{rg } (f) = n - \dim_K \text{Ker } f = n$ implica $\text{Ker } f = \{0\}$, luego f es inyectiva y recíprocamente (§ 140, corolario del teorema 2). Por otra parte, $\dim_K f(E) = \dim_K F$ implica $f(E) = F$ (§ 137, teorema 8).

Estas propiedades nos permiten enunciar el resultado siguiente:

COROLARIO 2. — *Si f es una aplicación lineal de* E *en* F, E *y* F *dos espacios vectoriales de igual dimensión finita, las propiedades siguientes son equivalentes:*

1. *f es biyectiva (es decir, es un isomorfismo de* E *sobre* F).
2. *f es inyectiva (es decir,* $\text{Ker } f = \{0\}$).
3. *f es suprayectiva (es decir,* $\text{Im } f = F$).

Estas propiedades se aplican evidentemente a todo endomorfismo de E de dimensión finita sobre K. Son falsas si E es de dimensión infinita (ver ej. 5 más abajo).

Estas propiedades muy particulares para una aplicación están relacionadas con las propiedades de las aplicaciones de un conjunto finito E en un conjunto finito F de igual cardinal (§ 31, corolario 5).

c) Sea $\{a_1, a_2, \ldots, a_n\}$ una base de E, f una aplicación de E en F de rango r; por tanto (teorema 7),

$$\dim_K \text{Ker } f = n - r$$

supongamos la base (a_i) escogida de manera que $\{a_1, a_2, \ldots, a_r\}$ sea una base

de *un suplementario del núcleo* de f y $\{a_{r+1}, \ldots, a_n\}$ una base de este núcleo, $\{f(a_1), \ldots, f(a_r)\}$ es una familia libre de $f(E)$, pues

$$\sum_{i=1}^{r} \alpha^i f(a_i) = 0 \Rightarrow f\left(\sum_{i=1}^{r} \alpha^i a_i\right) = 0$$

y $\sum_{i=1}^{r} \alpha^i a_i$ al pertenecer al núcleo de f y a su complementario es nulo, luego $\alpha^1 = \alpha^2 = \ldots = \alpha^r = 0$. Por otro lado, como para

$$i \in [r+1, n] \Rightarrow f(a_i) = 0$$

$\{f(a_1), \ldots, f(a_r)\}$ es una parte generatriz de $f(E)$, luego es una base de $f(E)$:

COROLARIO 3. — *Si f es una aplicación lineal de rango r de un espacio de dimensión finita n sobre K en un espacio vectorial F sobre K, toda base $\{a_1, a_2, \ldots, a_n\}$ de E tal que $\{a_{r+1}, \ldots, a_n\}$ sea una base de su núcleo es tal que $\{f(a_1), f(a_2), \ldots, f(a_r)\}$ es una base de $f(E)$.*

EJERCICIOS

1. Si E_1 y E_2 son dos subespacios vectoriales de dimensión finita de un espacio vectorial E se considera la aplicación f de $E_1 \times E_2$ en E definida por $f(x_1, x_2) = x_1 + x_2$ $(x_1 \in E_1, \ x_2 \in E_2)$.

a) Demostrar que f es lineal.

b) Demostrar que Ker f está descrito por $(x, -x)$, si x describe $E_1 \cap E_2$, deducir de lo anterior que Ker f es isomorfo a $E_1 \cap E_2$.

c) Demostrar que Im $f = E_1 + E_2$, deducir

$$\dim (E_1 + E_2) + \dim (E_1 \cap E_2) = \dim E_1 + \dim E_2$$

(utilizar § 136, ej. 4, y el teorema 7 de este §). Ver otra demostración de este resultado § 137, ej. 2. Si se considera la aplicación f de $E_1 \times E_2 \times \ldots \times E_p$ ($E_1, \ldots, E_p$ subespacios vectoriales de dimensión finita de E) en E definida por $f(x_1, \ldots, x_p) = x_1 + \ldots + x_p$ $(x_i \in E_i)$ demostrar que

$$\dim (E_1 + \ldots + E_p) \leq \dim E_1 + \ldots + \dim E_p.$$

2. Si E, F, G son tres espacios vectoriales de dimensión finita, se consideran las dos aplicaciones lineales $f\colon E \to F$ y $g\colon F \to G$; demostrar que

$$\dim (\operatorname{Im} f \cap \operatorname{Ker} g) = \operatorname{rg}(f) - \operatorname{rg}(g \circ f)$$
$$\sup (0, \operatorname{rg}(f) + \operatorname{rg}(g) - \dim F) \leq \operatorname{rg}(g \circ f) \leq \inf (\operatorname{rg}(f), \operatorname{rg}(g)).$$

(Utilizar § 137, ej. 2 o ej. 1, citado más arriba; § 137, ej. 3, y § 140, ej. 3.)

3. Si f y g son dos endomorfismos del espacio vectorial E, demostrar que

$$\operatorname{Ker}(g \circ f) = f^{-1}(\operatorname{Ker} g \cap \operatorname{Im} f).$$

4. Si f y g son dos endomorfismos de un espacio vectorial E de dimensión finita tal que $g \circ f = \mathrm{id}_E$, demostrar que f y g son dos automorfismos recíprocos de E.

5. Si E es un espacio vectorial de dimensión infinita numerable de base (a_i) $(i \in \mathbf{N}^*)$, se considera el endomorfismo f de E definido por

$$f(a_{2p+1}) = 0 \qquad f(a_{2p}) = a_p.$$

a) Demostrar que f es suprayectiva, no inyectiva.

b) Demostrar que existe un endomorfismo g inyectivo no suprayectivo tal que $f \circ g = \mathrm{id}_E$.

6. Si E y F son dos espacios vectoriales de dimensión finita, A y B dos subespacios vectoriales respectivos de E y de F tales que dim A + dim B = dim E, demostrar que existe al menos una aplicación lineal de E en F al que Ker f = A, Im f = B.

7. Si f y g son dos endomorfismos de un espacio vectorial E tales que $f \circ g = g \circ f$, demostrar que el núcleo o la imagen de uno de ellos es estable por el otro, es decir, por ejemplo, $f(\mathrm{Ker}\ g) \subset \mathrm{Ker}\ g$, $f(\mathrm{Im}\ g) \subset \mathrm{Im}\ g$.

8. Si E es un espacio vectorial de dimensión finita, E′ un subespacio de E, F un espacio vectorial, F′ un subespacio de F, se considera la aplicación lineal f: E → F; demostrar que

$$\dim f(E') = \dim E' - \dim (\mathrm{Ker}\ f \cap E')$$
$$\dim f^{-1}(F) = \dim (\mathrm{Im}\ f \cap F') + \dim E - \mathrm{rg}\ (f).$$

V. Operaciones algebraicas efectuadas sobre las aplicaciones lineales

144. Estructura de grupo abeliano y de espacio vectorial sobre K de $\mathcal{L}_K(E, F)$

Si f y g son dos elementos de $\mathcal{L}_K(E, F)$, definimos las aplicaciones $f + g$ y λf (λ elemento cualquiera de K) de E en F por

$$(\forall x \in E) \qquad (f + g)(x) = f(x) + g(x), \qquad (\lambda f)(x) = \lambda f(x)$$

a) f y g es lineal; en efecto, para todo x y todo y de E y para todo α de K

$$\begin{aligned}(f + g)(x + y) &= f(x + y) + g(x + y) && \text{(definición de } f + g)\\ &= [f(x) + f(y)] + [g(x) + g(y)] && \text{(1.ª propiedad de linealidad de } f \text{ y } g)\\ &= [f(x) + g(x)] + [f(y) + g(y)] && \text{(axioma } V_4 \text{ en F)}\\ &= (f + g)(x) + (f + g)(y) && \text{(definición de } f + g)\end{aligned}$$

$$\begin{aligned}(f + g)(\alpha x) &= f(\alpha x) + g(\alpha x) && \text{(definición de } f + g)\\ &= \alpha f(x) + \alpha g(x) && \text{(2.ª propiedad de linealidad de } f \text{ y } g)\\ &= \alpha[f(x) + g(x)] && \text{(axioma } V_8 \text{ en F)}\\ &= \alpha[(f + g)(x)] && \text{(definición de } f + g).\end{aligned}$$

Por tanto, $(f, g) \to f + g$ es una operación interna definida sobre $\mathcal{L}_K(E, F)$: su definición y las propiedades expresadas por los axiomas V_1 hasta el V_4 del espacio vectorial F (§ 125) demuestran que esta operación es asociativa y conmutativa; además, la aplicación ω de E en F definida por

$$(\forall x \in E) \qquad \omega(x) = 0_F$$

es evidentemente lineal, se le llama la *aplicación lineal nula* de E en F, para todo f de $\mathcal{L}_K(E, F)$ se tiene

$$f + \omega = \omega + f = f$$

$\mathcal{L}_K(E, F)$ posee un elemento neutro para la suma, si no da lugar a confusión se le puede representar (como 0_E, 0_F, 0_K) por el mismo signo 0 (§ 126, *c*). Por otro lado, se ve inmediatamente que la aplicación $-f$ definida para todo f de $\mathcal{L}_K(E, F)$ por

$$(\forall x \in E) \qquad (-f)(x) = -f(x)$$

es lineal y que

$$f + (-f) = (-f) + f = 0$$

por tanto: $\mathcal{L}_K(E, F)$ provisto de la suma interna $(f, g) \rightarrow f + g$ tiene una estructura de *grupo abeliano* (aditivo).

b) λf es lineal: en efecto, para todo x y todo y de E y todo α de K

$$\begin{aligned} (\lambda f)(x + y) &= \lambda f(x + y) && \text{(definición de } \lambda f) \\ &= \lambda[f(x) + f(y)] && \text{(1.ª propiedad de linealidad de } f) \\ &= \lambda f(x) + \lambda f(y) && \text{(axioma } V_8 \text{ en F)} \\ &= (\lambda f)(x) + (\lambda f)(y) && \text{(definición de } \lambda f) \end{aligned}$$

$$\begin{aligned} (\lambda f)(\alpha x) &= \lambda f(\alpha x) && \text{(definición de } \lambda f) \\ &= \lambda[\alpha f(x)] && \text{(2.ª propiedad de linealidad de } f) \\ &= (\lambda\alpha)f(x) && \text{(axioma } V_5 \text{ en F)} \\ &= (\alpha\lambda)f(x) && \text{(conmutatividad de la multiplicación en K)} \\ &= \alpha[\lambda f(x)] && \text{(axioma } V_5 \text{ en F)} \\ &= \alpha[(\lambda f)(x)] && \text{(definición de } \lambda f); \end{aligned}$$

en consecuencia, λf es un elemento de $\mathcal{L}_K(E, F)$; se observará que este hecho supone esencialmente la conmutatividad del cuerpo K (ver ej. 180). $(\lambda, f) \rightarrow \lambda f$ es, por tanto, una operación externa definida sobre $\mathcal{L}_K(E, F)$, si K es el dominio de operadores, dejamos al lector que demuestre que cualesquiera que sean f y g de $\mathcal{L}_K(E, F)$ y λ y μ de K, se tiene

$$\lambda(\mu f) = (\lambda\mu)f, \qquad 1f = f$$
$$(\lambda + \mu)f = \lambda f + \mu f, \qquad \lambda(f + g) = \lambda f + \lambda g.$$

TEOREMA. — *Si* E *y* F *son dos espacios vectoriales sobre el mismo cuerpo conmutativo* K, *el conjunto* $\mathcal{L}_K(E, F)$ *de las aplicaciones lineales de* E *en* F *provisto de la suma* $(f, g) \rightarrow f + g$ *y de la operación externa* $(\lambda, f) \rightarrow \lambda f$ (λ *elemento de* K) *tiene una estructura de espacio vectorial sobre* K.

Si dim E $= m$, dim F $= n$, se demuestra que dim $\mathcal{L}(E, F) = mn$ (ver § 154, *c*), y ej. 153, al final de este capítulo).

EJERCICIO

Si E, E′ son dos espacios vectoriales sobre K, isomorfos igualmente que F y F′, demostrar que $\mathcal{L}(E, F)$ es isomorfo a $\mathcal{L}(E', F')$. (Introducir los isomorfismos φ: $E \rightarrow E'$, ψ: $F \rightarrow F'$.)

145. Composición de aplicaciones lineales

a) Si f y g son dos elementos que pertenecen, respectivamente, a $\mathfrak{L}_K(E, F)$ y a $\mathfrak{L}_K(F, G)$, hemos visto en el § 140 (teorema 1) que $g \circ f$ es una aplicación lineal de E en G, luego un elemento de $\mathfrak{L}_K(E, G)$. Observemos que la aplicación definida por $(g, f) \to g \circ f$ *no es en general una ley de composición interna, ni una ley de composición externa*: es simplemente una aplicación de

$$\mathfrak{L}_K(F, G) \times \mathfrak{L}_K(E, F) \quad \text{en} \quad \mathfrak{L}_K(E, G).$$

Si f, g, h son tres aplicaciones lineales

$$E \xrightarrow{f} F \xrightarrow{g} G \xrightarrow{h} H$$

donde E, F, G, H son cuatro espacios vectoriales sobre el mismo cuerpo K, se tiene (§ 15)

$$(h \circ g) \circ f = h \circ (g \circ f)$$

por abuso de lenguaje se dice que la composición de las aplicaciones lineales es *asociativa.*

b) Sean tres espacios vectoriales E, F, G sobre el mismo cuerpo K, f_1 y f_2 dos elementos cualesquiera de $\mathfrak{L}(E, F)$ y g un elemento cualquiera $\mathfrak{L}(F, G)$, para todo x de E

$$\begin{aligned} [(f_1 + f_2) \circ g](x) &= (f_1 + f_2)[g(x)] && \text{(definición de } (f_1 + f_2) \circ g)\\ &= f_1[g(x)] + f_2[g(x)] && \text{(definición de } f_1 + f_2)\\ &= (f_1 \circ g)(x) + (f_2 \circ g)(x) && \text{(definición de } f_1 \circ g \text{ y } f_2 \circ g),\end{aligned}$$

por tanto,

$$(f_1 + f_2) \circ g = f_1 \circ g + f_2 \circ g.$$

Sea, por otra parte, f un elemento cualquiera de $\mathfrak{L}(E, F)$ y g_1 y g_2 dos elementos cualesquiera de $\mathfrak{L}(F, G)$, para todo x de E

$$\begin{aligned} [f \circ (g_1 + g_2)](x) &= f[(g_1 + g_2)(x)] && \text{(definición de } f \circ (g_1 + g_2))\\ &= f[g_1(x) + g_2(x)] && \text{(definición de } g_1 + g_2)\\ &= f[g_1(x)] + f[g_2(x)] && \text{(1.ª propiedad de linealidad de } f)\\ &= (f \circ g_1)(x) + (f \circ g_2)(x) && \text{(definición de } f \circ g_1 \text{ y } f \circ g_2),\end{aligned}$$

luego

$$f \circ (g_1 + g_2) = f \circ g_1 + f \circ g_2.$$

Por abuso de lenguaje, diremos que la *composición de aplicaciones lineales es distributiva por la derecha y por la izquierda respecto a la suma de aplicaciones lineales.*

OBSERVACION

Se ve que la distributividad por la derecha es independiente de la linealidad (ver capítulo 3, ej. 56).

146. Anillo $\mathcal{L}_K(E)$. Grupo $GL_K(E)$

a) Si hacemos ahora $E = F = G$, $(f, g) \to f + g$ y $(f, g) \to f \circ g$ son dos operaciones internas definidas sobre $\mathcal{L}(E)$.

Para la adición $\mathcal{L}(E)$ es un grupo abeliano (§ 144); por otra parte, $(f, g) \to f \circ g$ es asociativa y distributiva por la izquierda y por la derecha respecto a la adición (§ 145). Finalmente la aplicación idéntica de E en E definida como se sabe (§ 12, *c*) por

$$(\forall x \in E) \qquad id_E(x) = x$$

es evidentemente lineal y para todo endomorfismo del espacio vectorial E

$$f \circ id_E = id_E \circ f = f$$

$\mathcal{L}(E)$ posee, pues, un elemento neutro para la multiplicación, lo representaremos generalmente por id_E, algunos autores lo representan 1_E, de donde:

TEOREMA. — *Si* E *es un espacio vectorial sobre el cuerpo conmutativo* K, *el conjunto* $\mathcal{L}(E)$ *de los endomorfismos de* E *provisto de las operaciones internas* $(f, g) \to f + g$ *y* $(f, g) \to f \circ g$ *tiene una estructura de anillo unitario.*

Como los ejemplos simples lo demuestran, este anillo es en general ***no conmutativo*** (ej. 1) y está provisto de *divisores de cero* (ej. 2 y 3).

b) El conjunto de los *automorfismos* del espacio vectorial E sobre K es una parte de $\mathcal{L}(E)$; si f y g son automorfismos de E, también lo es $f \circ g$ (§ 140, corolario del teorema 1); por otra parte, id_E es un automorfismo de E; finalmente si f es un automorfismo de E, igualmente lo es f^{-1} (§ 140, teorema 3), de donde:

TEOREMA Y DEFINICIÓN. — *El conjunto de los automorfismos de un espacio vectorial* E *sobre un cuerpo conmutativo* K *es, para la composición de las aplicaciones, un grupo, llamado* grupo lineal *de* E *y representado* $GL_K(E)$.

OBSERVACIONES

1. Este grupo $GL_K(E)$ es precisamente el grupo de los elementos inversibles del anillo $\mathcal{L}_K(E)$ (§ 98, t. 1).

2. Se demuestra fácilmente (ej. 4, más abajo) que si los dos espacios vectoriales E y E′ sobre K son isomorfos, lo mismo ocurre con los grupos $GL_K(E)$ y $GL_K(E')$; en particular, cualquiera que sea E de dimensión n sobre K el grupo lineal de E es isomorfo al grupo lineal de K^n: esta estructura de grupo, único, está, por tanto, completamente determinado por K y n, se le representa $GL_n(K)$.

EJERCICIOS

1. Si E es el espacio vectorial sobre **R** de las funciones numéricas $t \to x(t)$ de [0, 1] en **R**, indefinidamente derivables (por tanto, continuas), se considera los dos endomorfismos de E (§ 139, ej. 4 y 5)

$$x \to f(x) = x', \qquad x \to g(x) \qquad \text{con} \qquad [g(x)](t) = \int_0^t x(u)\, du.$$

Demostrar que si $x(0) \neq 0$: $(g \circ f)(x) \neq (f \circ g)(x)$.

2. Dados dos endomorfismos f y g de E tales que Im $f \subset$ Ker g, demostrar que $g \circ f = 0$. Deducir que si f no es divisor de cero por la derecha en el anillo $\mathcal{L}(E)$, f es suprayectiva y que si f no es divisor de cero por la izquierda, f es inyectiva. Demostrar que todo endomorfismo no nulo de E no divisor de cero ni por la derecha ni por la izquierda es un automorfismo.

3. Si E es un espacio vectorial de dimensión n sobre K, ¿qué desigualdad verifican los rangos de dos endomorfismos f y g tales que $g \circ f = 0$?

¿Hay endomorfismos f y g tales que $g \circ f = f \circ g = 0$? (Utilizar el ejercicio 6 del § 143.)

4. Si E y E′ son dos espacios vectoriales sobre K isomorfos, demostrar que $\mathcal{L}_K(E)$ y $\mathcal{L}_K(E')$ son isomorfos y que también lo son $\mathbf{GL}_K(E)$ y $\mathbf{GL}_K(E')$ (introducir el isomorfismo φ: $E \to E'$).

5. Sea $\mathcal{H}_K(E)$ el conjunto de las homotecias de E, descrito por f_λ ($\lambda \in K$), tal que $f_\lambda(x) = \lambda x$ para todo x de E (§ 125, ej. 8). Demostrar que $f_\lambda \circ g = g \circ f_\lambda$ para todo endomorfismo g de E (ver un recíproco en ej. 155 al final del capítulo). Demostrar que $\mathcal{H}_K(E)$ es un subanillo de $\mathcal{L}(E)$ que tiene una estructura de cuerpo isomorfo a la de K.

6. Dados f y g que pertenecen a $\mathcal{L}(E)$ y siendo E de dimensión finita, demostrar que $f \circ g = id_E$ implica $g = f^{-1}$.

147. Noción de álgebra sobre un cuerpo conmutativo K. Algebra $\mathcal{L}(E)$

a) Si E es un espacio vectorial sobre el cuerpo conmutativo K, $\mathcal{L}(E)$ está provisto:

(1) de una suma interna

$$(f, g) \to f + g.$$

(2) de una multiplicación externa

$$(\lambda, f) \to \lambda f.$$

(3) de una multiplicación interna

$$(f, g) \to f \circ g.$$

Para (1) y (2): $\mathcal{L}(E)$ tiene una estructura de *espacio vectorial* sobre K, para (1) y (3): $\mathcal{L}(E)$ tiene una estructura de *anillo*. Además, se verifica fácilmente que para todo λ de K y todo par de endomorfismos de $\mathcal{L}(E)$, se tiene

$$\lambda(f \circ g) = (\lambda f) \circ g = f \circ (\lambda g)$$

estas propiedades nos llevan a establecer la definición siguiente:

DEFINICIÓN. — *Dado un cuerpo conmutativo* K *y un conjunto* E *provisto de una adición, de una multiplicación interna y de una multiplicación externa cuyo dominio de operadores es* K, *se dice que* E *tiene una* estructura de álgebra *sobre* K *si:*

1. E *tiene una estructura de espacio vectorial sobre* K *para la adición interna y la multiplicación externa.*

2. E *tiene una estructura de anillo para la adición y la multiplicación internas.*

3. *Para todo* λ *de* K *y todo par* (x, y) *de elementos de* E, *se tiene*

$$\lambda(xy) = (\lambda x)y = x(\lambda y).$$

EJEMPLOS

1. Si E es un espacio vectorial sobre el cuerpo conmutativo K, $\mathfrak{L}(E)$ provisto de las operaciones $f+g$, λf, $f \circ g$ es una álgebra sobre K.

2. Todo cuerpo conmutativo K es una álgebra sobre sí mismo.

3. El *cuerpo* de los complejos **C** que es también un *espacio vectorial* sobre **R** puede ser considerado como una *álgebra* sobre **R**.

4. K^n descrito por $x = (x_1, x_2, ..., x_n)$ y provisto de las operaciones

$$\begin{aligned} x + y &= (x_1 + y_1, ..., x_i + y_i, ..., x_n + y_n) \\ xy &= (x_1y_1, ..., x_iy_i, ..., x_ny_n) \\ \lambda x &= (\lambda x_1, ..., \lambda x_i, ..., \lambda x_n) \end{aligned}$$

(ver ej. 92, fin del capítulo 5) es una álgebra sobre K, si $\{e_1, e_2, ..., e_n\}$ es la base canónica de K^n, hallar la tabla de multiplicar de los elementos de esta base.

5. $\mathfrak{F}(\mathbf{R}, \mathbf{R})$ provisto de las operaciones $f+g$, fg, λf (ver § 90, ej. 4, y § 125, ej. 5) es una álgebra sobre **R**.

6. Veremos (capítulo 11) que el conjunto $\mathfrak{F}$ de los polinomios en x con coeficientes reales provisto de las operaciones $A + B$, λA, AB es una álgebra sobre **R**.

b) Dada una álgebra E sobre el cuerpo conmutativo K, una base del espacio vectorial E se le llamará también *base del álgebra* E; igualmente si el espacio vectorial E es de dimensión n sobre K, diremos que el álgebra E es de *dimensión* n sobre K.

Por otra parte, se comprobará fácilmente que todo subanillo F de E, estable para la multiplicación externa tiene una estructura de álgebra sobre K para las tres operaciones inducidas sobre F por las operaciones definidas sobre E; se dice que F es una *subálgebra* de E. Igualmente se llamará *ideal* del álgebra E, todo ideal del anillo E estable para la operación externa.

Se ve inmediatamente que la intersección de una familia cualquiera de subálgebras del álgebra de E sobre K es una subálgebra de E; dada una parte no vacía X del álgebra E, la intersección de todas las subálgebras de E conteniendo X se llama *subálgebra engendrada por* X (ver §§ 93 y 129).

De acuerdo con las definiciones generales (§ 69), una aplicación f de una álgebra E en una álgebra E′, las dos sobre el mismo cuerpo conmutativo K, tal que para todo x, todo y de E y todo α de K

$$f(x+y) = f(x) + f(y), \qquad f(xy) = f(x)f(y), \qquad f(\alpha x) = \alpha f(x)$$

se llamará un *homomorfismo* de álgebras; se definirá igualmente un *isomorfismo* de álgebras y un *endomorfismo* o un *automorfismo* de una álgebra. Se verificará sin dificultad que los teoremas análogos a los teoremas 1, 2 y 3 (§§ 96 y 140) son exactos.

EJEMPLOS Y EJERCICIOS

7. La subálgebra de E, álgebra sobre el cuerpo conmutativo K, engendrado por e, elemento unidad de K es isomorfo a K considerado como álgebra sobre sí mismo: se le identifica a menudo con K; por tanto, $K \subset E$.

8. El conjunto $\mathcal{H}(E)$ de las homotecias del espacio vectorial E, sobre el cuerpo conmutativo K (§ 146, ej. 5) es una subálgebra de la álgebra $\mathfrak{L}(E)$ sobre K, isomorfa a K considerado como álgebra sobre sí mismo (ej. 7 más arriba):

9. El conjunto de las funciones *polinomios* reales de variable real (ver capítulo 11) es una subálgebra del álgebra $\mathfrak{F}(\mathbf{R}, \mathbf{R})$ sobre **R** (ej. 5 anterior) esta subálgebra es isomorfa al álgebra $\mathfrak{F}$ de los polinomios en x con coeficientes reales (ej. 6 anterior).

10. Si E es el conjunto de las funciones reales de variable real *indefinidamente* derivables para todo t de **R**, demostrar que E es una subálgebra de $\mathfrak{F}(\mathbf{R}, \mathbf{R})$ (ej. 5 anterior). Entre las aplicaciones siguientes de E en $F = f(E)$ es una álgebra sobre **R** las que son homomorfismos de álgebras:

a) $f(x) = x(t_0)$ (valor de la función x para t_0 fijo).
b) $f(x) = x'$ (función derivada de la función x).
c) $f(x) = x'(t_0)$ (valor de x' para t_0 fijo)
d) $f(x) = X$ (X primitiva de x, tal que $X(0) = 0$).
e) $f(x) = P_n(t)$ (parte regular de orden n de *un desarrollo limitado* de la función x alrededor de cero).
f) $f(x) = l$ (límite de la función x para t_0 fijo).

(Ver curso de *Análisis*.)

c) Factorización y linealización en una álgebra

Sea x un elemento de una álgebra E sobre un cuerpo conmutativo K, que se puede escribir en las dos formas

(1) $x = a_1 \dots a_i \dots a_p$ $\quad (a_i \in E, \quad i = 1, \dots, p)$
(2) $x = \lambda^1 b_1 + \dots + \lambda^j b_j + \dots + \lambda^q b_q$ $\quad (b_j \in E, \quad \lambda_j \in K, \quad j = 1, 2, \dots, q)$

los pasos de la forma (1) a la forma (2) y de la forma (2) a la forma (1) se llaman, respectivamente, *linealización* y *factorización*: se halla así una generalización de las transformaciones de "producto en suma" y de "suma en producto" de la trigonometría elemental.

Si es un anillo de integridad y $\bar{E}$ su cuerpo de fracciones, la *factorización* es útil para reducir las fracciones y de una manera general para calcular en E, para resolver las ecuaciones, etc.

La *linealización* se utiliza cuando se quiere calcular $f(x)$, conociendo $f(b_1) \dots f(b_q)$, si f es *solamente* un homomorfismo de espacios vectoriales sobre K y no un homomorfismo de álgebras sobre K, lo que se produce muy a menudo en la práctica (ver ej. 10 anterior).

VI. Formas lineales. Dualidad

148. Formas lineales. Dual de un espacio vectorial

Sea E un espacio vectorial sobre el cuerpo conmutativo K, podemos considerar K como espacio vectorial sobre sí mismo (§ 125, ej. 3), $\mathfrak{L}(E, K)$ es entonces un espacio vectorial sobre K; podemos, pues, dar las definiciones siguientes:

DEFINICIÓN. — *Dado un espacio vectorial* E *sobre el cuerpo conmutativo* K, *toda aplicación lineal de* E *en* K *se llama* forma lineal *definida sobre* E.

El conjunto $\mathcal{L}(E, K)$ *de las formas lineales definidas sobre* E, *espacio vectorial sobre* K, *se llama* dual *de* E *y se representa* E^*(30). *El dual de* E^* *se llama el* bidual *de* E *y se representa* E^{**}.

Si f es una forma lineal definida sobre E, se tendrá para todo x, todo y de E y todo α de K

$$f(x + y) = f(x) + f(y), \qquad f(\alpha x) = \alpha f(x).$$

OBSERVACION

Se representarán a menudo x^*, y^* ... los elementos de E^*, sin que esta notación implique forzosamente la existencia de una relación funcional entre x de E y x^* de E^*. Todas las nociones y operaciones relativas a los elementos de un espacio vectorial se aplican, por tanto, a las formas lineales, se hablará de formas lineales independientes, de suma de dos formas, etc.

EJEMPLOS

1. Sea $x = (\alpha^1, \ldots, \alpha^i, \ldots \alpha^n)$ de K^n expresado en su base canónica (a_i) (§ 134, ej. 1), a toda familia (λ_i) $(1 \leq i \leq n)$ de elementos de K, la fórmula

$$f(x) = \lambda_1\alpha^1 + \ldots + \lambda_i\alpha^i + \ldots + \lambda_n\alpha^n$$

define una aplicación f de K^n en K, que es una forma *lineal* definida sobre K^n. Se dice algunas veces que $\lambda_1\alpha^1 + \ldots + \lambda_i\alpha^i + \ldots + \lambda_n\alpha^n$ es una forma lineal: es el abuso del lenguaje clásico que confunde la aplicación f y su valor $f(x)$ para el valor x de la variable (§ 11, observación). Se tiene naturalmente $f(a_i) = \lambda_i$.

2. E de dimensión n sobre K si tiene (a_i) por base, para todo x de E, se tiene $x = \sum_{i=1}^{i=n} \alpha^i a_i$ la aplicación f^i de E en K definida por

$$x \rightarrow f^i(x) = \alpha_i$$

es una forma lineal definida sobre E llamada la i-ésima *forma coordenada.*

3. Si E es el espacio vectorial sobre **R** de las funciones numéricas aplicando $[0, 1]$ en **R** y derivables para t_0 fijo de $[0, 1]$, la aplicación de E en **R** definida por $x = x'(t_0)$ es una forma lineal sobre E (ver curso de Análisis).

4. Si E es el espacio vectorial sobre **R** de las funciones numéricas aplicando $[0, 1]$ en **R** y continuas sobre $[0, 1]$, la aplicación de E en **R** definida por $x \rightarrow \int_0^1 x(t)\,dt$ es una forma lineal definida sobre E.

149. Forma bilineal canónica definida sobre E × E*

a) Sean dos espacios vectoriales E_1, E_2 sobre el mismo cuerpo conmutativo K, consideremos una aplicación F de $E_1 \times E_2$ en K tal que las aplicaciones parciales (§ 12, *d*) $F(., x_2)$ de E_1 en K y $F(x_1, .)$ de E_2 en K sean

(30) Para un cuerpo K designamos $K^* = K - \{0\}$; si se considera K como espacio vectorial sobre sí mismo el contexto indicará si K^* designa $K - \{0\}$ o el dual del espacio vectorial K.

lineales, obtendremos, por tanto, idénticamente (es decir, cualesquiera que sean x_1, y_1 de E_1, x_2, y_2 de E_2, λ de K)

$$F(x_1 + y_1, x_2) = F(x_1, x_2) + F(y_1, x_2), \qquad F(\lambda x_1, x_2) = \lambda F(x_1, x_2)$$
$$F(x_1, x_2 + y_2) = F(x_1, x_2) + F(x_1, y_2), \qquad F(x_1, \lambda x_2) = \lambda F(x_1, x_2).$$

Una aplicación de $E_1 \times E_2$ en K que verifica estas identidades se llama *forma bilineal definida sobre* $E_1 \times E_2$. La aplicación parcial $F(\,.\,, x_2)$ *asociada* al elemento x_2 de E_2, es una *forma lineal definida sobre* E_1, igual $F(x_1, \,.\,)$, forma lineal definida sobre E_2 es asociada al elemento x_1 de E_1.

EJEMPLO

En $\mathbf{R}^3$ si $x = (\alpha, \beta, \gamma)$, $x' = (\alpha', \beta', \gamma')$ la aplicación $(x, x') \rightarrow \alpha\alpha' + \beta\beta' + \gamma\gamma'$ es una forma bilineal definida sobre $\mathbf{R}^3 \times \mathbf{R}^3$.

b) Sea E un espacio vectorial sobre K, E^* su dual, si x es un elemento cualquiera de E, f un elemento cualquiera del dual de E, consideremos la aplicación de $E \times E^*$ en K definida por

$$(x, f) \rightarrow \langle x, f \rangle = f(x)$$

utilizando la linealidad de f, las propiedades del espacio vectorial $\mathfrak{L}(E, K) = E^*$ y naturalmente la definición de $\langle x, f \rangle$, tendremos las identidades siguientes

(1) $\langle x + y, f \rangle = f(x + y) = f(x) + f(y) = \langle x, f \rangle + \langle y, f \rangle$

(1.ª propiedad de linealidad de f)

(1′) $\langle \lambda x, f \rangle = f(\lambda x) = \lambda f(x) = \lambda \langle x, f \rangle$

(2.ª propiedad de linealidad de f)

(2) $\langle x, f + g \rangle = (f + g)(x) = f(x) + g(x) = \langle x, f \rangle + \langle x, g \rangle$

(definición de la adición en $E^* = \mathfrak{L}(E, K)$)

(2′) $\langle x, \lambda f \rangle = (\lambda f)(x) = \lambda f(x) = \lambda \langle x, f \rangle$

(definición de λf en $E^* = \mathfrak{L}(E, K)$)

(1) y (1′) traducen el hecho de que la aplicación parcial $\langle\,.\,, f \rangle$ de E en K, que no es otra que f, es una *forma lineal definida sobre* E.

(2) y (2′) traducen el hecho que la aplicación parcial $\langle x, \,.\, \rangle$ de E^* en K es una *forma lineal definida sobre* E^*; por tanto, la aplicación considerada es una forma bilineal:

TEOREMA Y DEFINICIÓN. — *Dado un espacio vectorial* E *sobre un cuerpo conmutativo* K, E^* *su dual, la aplicación de* $E \times E^*$ *en* K *definida por*

$$(x, f) \rightarrow \langle x, f \rangle = f(x)$$

es una forma bilineal: se llama forma bilineal canónica *definida sobre* $E \times E^*$[31].

(31) El símbolo <x, f> se llama algunas veces *corchete de dualidad*.

c) Estudiemos con más detalle la aplicación parcial $\langle x, . \rangle$ definida por

$$f \to \langle x, f \rangle = f(x)$$

es una forma lineal definida sobre E*; es, por tanto, un elemento del bidual E** de E. Por otra parte, esta forma lineal está asociada al elemento x de E, la representaremos $\tilde{x}$ (que se lee "x tilda"). Su valor para f, o sea, $\tilde{x}(f)$, es por definición $\langle x, f \rangle$, pero $(f, \tilde{x}) \to \tilde{x}(f)$ es la forma bilineal canónica definida sobre $E^* \times E^{**}$, de donde $\tilde{x}(f) = \langle f, \tilde{x} \rangle$ y, por consiguiente,

$$\langle f, \tilde{x} \rangle = \langle x, f \rangle \tag{3}$$

y esto cualquiera que sea x de E y f de E*. *La fórmula* (3) *permite, por tanto, hacer corresponder a todo x de* E *un elemento único $\tilde{x}$ de* E**.

Estudiemos esta aplicación $x \to \tilde{x}$ de E en E**, primero vemos que es *canónica* en el sentido que depende únicamente de la estructura de espacio vectorial de E, seguidamente vemos que es *lineal*; en efecto, utilizando las propiedades de la forma bilineal canónica y la fórmula (3) tenemos idénticamente

$$\left(\widetilde{x+y}\right)(f) = \left\langle f, \widetilde{x+y} \right\rangle = \langle x+y, f \rangle = \langle x, f \rangle + \langle y, f \rangle$$
$$= \langle f, \tilde{x} \rangle + \langle f, \tilde{y} \rangle = \tilde{x}(f) + \tilde{y}(f)$$
$$\left(\widetilde{\lambda x}\right)(f) = \left\langle f, \widetilde{\lambda x} \right\rangle = \langle \lambda x, f \rangle = \lambda \langle x, f \rangle$$
$$= \lambda \langle f, \tilde{x} \rangle = \lambda \tilde{x}(f),$$

puesto que estas igualdades se verifican para todo f de E*, tenemos

$$\widetilde{x+y} = \tilde{x} + \tilde{y}, \qquad (\widetilde{\lambda x}) = \lambda \tilde{x}$$

cualesquiera que sean x e y de E y λ de K.

OBSERVACION SOBRE LAS NOTACIONES

Para mayor facilidad hemos designado por x y por f los elementos que describen, respectivamente, E y E*, para señalar que f describe E* podemos representarlo por una letra con asterisco —por ejemplo, x^* o y^*— recordando que la notación x^* no supone forzosamente la existencia de una relación funcional entre x y x^* (lo que no es el caso para la notación $\tilde{x}$, que indica un elemento de E** asociado a x). Podemos así enunciar los resultados precedentes de la siguiente manera:

Teorema y definición. — *Dado un espacio vectorial* E *descrito por x, su dual* E* *descrito por x^* y la forma bilineal canónica $(x, x^*) \to \langle x, x^* \rangle$ definida sobre* $E \times E^*$, *la aplicación parcial definida por $x^* \to \langle x, x^* \rangle$ es un elemento $\tilde{x}$ del bidual de* E, *está definida por*

$$(\forall x \in E) \quad (\forall x^* \in E^*) \quad \langle x^*, \tilde{x} \rangle = \langle x, x^* \rangle. \tag{4}$$

La aplicación $x \to \tilde{x}$ de E *en* E** *es lineal, se llama* homomorfismo canónico *de* E *en* E**.

Recordemos una vez más que $\langle x, x^* \rangle = x^*(x)$.

150. Caso en que E es de dimensión finita. Bases duales

a) Supongamos E de dimensión n sobre K y $\{a_1, a_2, \ldots, a_n\}$ una base de E; para todo x de E

(1) $$x = \sum_{i=1}^{n} \alpha^i a_i$$

sea x^* una forma lineal definida sobre E, es decir, un elemento del dual, tendremos

(2) $$x^*(x) = \langle x, x^* \rangle = \sum_{i=1}^{n} \alpha^i x^*(a_i) = \sum_{i=1}^{n} \alpha^i \alpha_i^*$$

poniendo $\alpha_i^* = x^*(a_i)$ (que es un elemento de K) encontramos así de nuevo un caso particular del teorema del § 143: *Toda forma lineal definida sobre E de dimensión n está determinada al dar los n valores que toma sobre los valores de una base de* E.

En particular, podemos definir n formas coordenadas (§ 148, ej. 2), que representaremos por $a^{*1}, \ldots, a^{*i}, \ldots, a^{*n}$, definidas por

(3) $$a^{*i}(x) = \alpha^i$$

o, lo que equivale según el resultado recordado antes, por las igualdades

$$i \neq j \qquad a^{*j}(a_i) = \langle a_i, a^{*j} \rangle = 0, \qquad a^{*i}(a_i) = \langle a_i, a^{*i} \rangle = 1$$

es decir, utilizando el símbolo de KRONECKER (§ 133, ej. 1)

(4) $$a^{*j}(a_i) = \langle a_i, a^{*j} \rangle = \delta_i^j$$

gracias a (3) la fórmula (2) se convierte en

(2′) $$x^*(x) = \langle x, x^* \rangle = \sum_{i=1}^{n} \alpha_i^* a^{*i}(x)$$

de donde, siendo (2′) verdadera para todo x de E,

(5) $$x^* = \sum_{i=1}^{n} \alpha_i^* a^{*i}$$

como esta fórmula es cierta para todo x^* de E*, vemos que las n formas coordenadas $a^{*1}, \ldots, a^{*n}$ engendran E*. Vamos a ver que estas formas describen una base de E*: nos basta demostrar que son linealmente independientes (§ 136, corolario del teorema 7), pongamos

$$\sum_{i=1}^{n} \lambda_i a^{*i} = 0 \qquad \text{(igualdad en E*)}$$

esto significa que para todo x de E

$$\sum_{i=1}^{n} \lambda_i a^{*i}(x) = 0 \qquad \text{(igualdad en K)}$$

en particular, para $x = a_j$ obtendremos teniendo en cuenta (4)

$$\sum_{i=1}^{n} \lambda_i a^{*i}(a_j) = \sum_{i=1}^{n} \lambda_i \delta_i^j = \lambda_j = 0$$

por tanto, las n formas coordenadas son completamente independientes; podemos resumir todos estos resultados:

TEOREMA Y DEFINICIÓN.—*Si* E *es un espacio vectorial de dimensión* n *sobre* K, *expresado en una base* $B = \{a_1, \ldots, a_n\}$ *las* n *formas coordenadas* a^{*j} $(1 \leq j \leq n)$ *definidas por*

$$a^{*j}(a_i) = \langle a_i, a^{*j} \rangle = \delta_i^j$$

describen una base B* *del dual* E* *de* E *llamada* base dual *de la base* B *de* E. *Por* otra parte,

$$\dim_K E^* = \dim_K E = n$$

y si

$$x = \sum_{i=1}^{n} \alpha^i a_i, \qquad x^* = \sum_{j=1}^{n} \alpha_j^* a^{*j}$$

se tiene

$$x^*(x) = \langle x, x^* \rangle = \sum_{i=1}^{n} \alpha^i \alpha_i^*.$$

OBSERVACION

En este caso (E de dimensión finita) E y su dual E*, al tener la misma dimensión sobre K, son isomorfos: hay en general una infinidad de isomorfismos φ de E sobre E*. (Si φ es uno de ellos y u un automorfismo de E, $\varphi \circ u$ es un isomorfismo φ' en general distinto de φ.)

Se demuestra (ver ej. 171, al final del capítulo) que no existe isomorfismo canónico de E sobre E* si $n > 1$, salvo si $n = 2$ y $K = \mathbf{Z}/2\mathbf{Z}$.

b) En el caso en que E es de *dimensión finita,* podemos precisar el teorema del § 149, *c*) demostrando que el homomorfismo canónico $x \longrightarrow \tilde{x}$ de E en E** es un isomorfismo. Según el teorema anterior, como $E^{**} = (E^*)^*$; $\dim E^{**} = \dim E^* = \dim E$; nos basta, por tanto, de demostrar que la aplicación de E en E**, definida por $x \longrightarrow \tilde{x}$, es inyectiva (§ 143, corolario 2).

Para todo x^* de E*

$$\tilde{x} = \tilde{y} \Rightarrow \langle x^*, \tilde{x} \rangle = \langle x^*, \tilde{y} \rangle \Leftrightarrow \langle x, x^* \rangle = \langle y, x^* \rangle$$

$$\tilde{x} = \tilde{y} \Rightarrow \langle x - y, x^* \rangle = 0 \Rightarrow x^*(x - y) = 0$$

en particular para cada elemento a^{*i} de la base de E* dual de la base (a_i) $(1 \leq i \leq n)$ de E se tiene

$$a^{*i}(x - y) = \alpha^i - \beta^i = 0$$

donde α^i y β^i son las i-ésimas coordenadas respectivas de x e y, pues $x = y$, de donde:

TEOREMA. — *Si* E *es de dimensión finita sobre* K, *la aplicación canónica de* E *en* E**, $x \to \tilde{x}$ *definida por*

$$\langle x^*, \tilde{x} \rangle = \langle x, x^* \rangle$$

es un isomorfismo de espacios vectoriales.

Este isomorfismo es *canónico* (la consideración de la base B sólo ha servido para demostrar que el homomorfismo canónico $x \to \tilde{x}$ era biyectivo); como no hay ningún interés para considerar E** por sí mismo, *identificaremos* E con E** designando x y $\tilde{x}$ por el mismo símbolo x; gracias a esta identificación E y E* son *duales uno del otro,* la base B** identificada a B es dual de la base B* de E*, se dice que B y B* son dos *bases duales.*

Por otra parte, x^* *elemento* de E* es una *forma lineal definida sobre* E; igualmente la aplicación $x^* \to \langle x, x^* \rangle$, o sea, $\tilde{x}$, identificada a x es un *elemento* de E y es una *forma lineal definida sobre* E*; hay también *reciprocidad perfecta* entre E y E*: cada elemento de uno de estos espacios vectoriales es una forma lineal definida sobre el otro.

OBSERVACION

En el caso en que E es de *dimensión infinita* sobre K, se demuestra que la aplicación $x \to \tilde{x}$ es siempre inyectiva, pero nunca sobreyectiva (ver ej. 172, fin del capítulo).

EJERCICIOS

1. Determinar la base de $(K^n)^*$ dual de la base canónica de K^n (§ 134, ej. 1).
2. Si E, de dimensión finita está expresado en una base B, demostrar que en E** la base dual de la base B* de E*, también dual de B, es la imagen de B en el isomorfismo canónico $x \to \tilde{x}$ de E sobre E**.

191. Ortogonalidad

a) DEFINICIÓN. — *Se dice que un elemento x de un espacio vectorial* E *y un elemento x^* del dual* E* *de* E *son* ortogonales *si y sólo si* $\langle x, x^* \rangle = 0$. *Se dice también que x es ortogonal a x^* o que x^* es ortogonal a x.*

Se dice que una parte A *de* E *y una parte* A* *de* E* *son ortogonales si y sólo si todo elemento de* A *es ortogonal a todo elemento de* A*.

Si x es ortogonal a x^* y a y^*, será ortogonal a λx^* y a $x^* + y^*$; por tanto, si x es ortogonal a una parte A^* de E^* es ortogonal al subespacio de E^* engendrado por A^*; en particular, tenemos el resultado siguiente:

TEOREMA Y DEFINICIÓN. — *Si* F *es un subespacio vectorial de* E, *el conjunto de los elementos de* E^* *ortogonales a todos los elementos de* F *es un subespacio vectorial de* E^* *llamado el* subespacio ortogonal *de* F *en* E^* *y se representa por* $F^\perp$.

Se definirá igualmente el subespacio vectorial ortogonal $(F^*)^\perp$ de F^* en E; por ejemplo,

$$\{0_E\}^\perp = E^*, \qquad E^\perp = \{0_{E^*}\}.$$

OBSERVACION

Si G' es un subespacio propio de $F^\perp$ es *un* subespacio de E^* ortogonal a F, pero no es *el* subespacio ortogonal a F; para evitar esta posibilidad de ambigüedad algunos autores llaman a $F^\perp$ el *subespacio totalmente ortogonal a* F. (Nos encontramos en la misma situación que en el espacio métrico $\mathbf{R}^3$, cuando decimos que un plano P' y una recta D' los dos pasando por O son ortogonales a una recta D pasando por O, tenemos $D' \subset P'$, $D' \neq P'$, se podría decir que P' es totalmente ortogonal a D, ver libro III Geometría.)

b) Supongamos ahora dim $E = n$. Sea F un subespacio vectorial de E de dimensión $0 < m < n$, consideremos una base (a_i) $(1 \leq i \leq m)$ de F que completada da una base (a_i) $(1 \leq i \leq n)$ de E (§ 135, corolario del teorema 4), o sea, (a^{*i}) $(1 \leq i \leq n)$ la base dual de E^*, para que x^* pertenezca a $F^\perp$ es necesario y suficiente que

$$1 \leq i \leq m \Rightarrow \langle a_i, x^* \rangle = 0.$$

Pongamos

$$x^* = \sum_{j=1}^{n} \alpha_j^* a^{*j}$$

se tendrá

$$1 \leq i \leq m \qquad \left\langle a_i, \sum_{j=1}^{n} \alpha_j^* a^{*j} \right\rangle = \sum_{j=1}^{n} \alpha_j^* \langle a_i, a^{*j} \rangle = \sum_{j=1}^{n} \alpha_j^* \delta_i^j = \alpha_i^* = 0$$

por tanto, $F^\perp$ es el conjunto de los vectores de E^* de la forma

$$x^* = \alpha_{m+1}^* a^{*m+1} + \ldots + \alpha_n^* a^{*n}$$

$F^\perp$ es, por tanto, el subespacio de E^* engendrado por $a^{*m+1}, \ldots, a^{*n}$, luego es de dimensión $n - m = \dim E - \dim F = \operatorname{codim}_E F$.

Consideremos ahora el ortogonal $(F^\perp)^\perp$ de E en $F^\perp$ (suponiendo identificados E y E^{**}; ver § 150, *b*). Todo vector de F es por definición ortogonal a $F^\perp$; por tanto,

$$F \subset (F^\perp)^\perp;$$

por otra parte,

$$\dim (F^{\perp})^{\perp} = \dim E^* - \dim F^{\perp} = n - (n - m) = m = \dim F;$$

por tanto (§ 137, teorema 8),

$$(F^{\perp})^{\perp} = F,$$

como para $m = n$: $(E)^{\perp} = \{0_{E^*}\}$ y $\{0_{E^*}\}^{\perp} = E$ tenemos:

TEOREMA. — *Dado un espacio vectorial* E *de dimensión* n *sobre* K, *el ortogonal* $F^{\perp}$ *de un subespacio vectorial* F *de* E, *de dimensión* $m \leq n$ *es un subespacio vectorial de* E^* *de dimensión* $n - m$. *El ortogonal de* $F^{\perp}$ *es* F.

Dicho de otra manera la relación entre F (de E) y $G^* = F^{\perp}$ (de E^*) es *simétrica.*

EJEMPLOS

1. Si $E = K^n$ está expresado en la base canónica (§ 134, ej. 1) y E^* en la base dual (§ 150, ej. 1), sea $u = (u_1, u_2, \ldots, u_n)$ un elemento de E^*; el ortogonal de la recta (§ 137) engendrada por u (por tanto, de dimensión 1) es de dimensión $n - 1$, es un hiperplano (§ 137) de E; está descrito por $x = (x^1, x^2, \ldots, x^n)$ de E tal que

$$u(x) = \langle x, u \rangle = u_1x^1 + u_2x^2 + \ldots + u_nx^n = 0$$

se le llama el *hiperplano ortogonal a u,* la relación precedente es la *ecuación cartesiana* del hiperplano.

2. Con las mismas notaciones si F^* está engendrado por $u^j = (u^j_1, \ldots, u^j_i, \ldots, u^j_n)$ $(1 \leq j \leq p)$ el ortogonal $(F^*)^{\perp}$ es un subespacio de E descrito por $x = (x^1, x^2, \ldots, x^n)$ de E tal que

$$\begin{cases} u^1(x) = \langle x, u^1 \rangle = u^1_1x^1 + \ldots + u^1_ix^i + \ldots + u^1_nx^n = 0 \\ \cdot \\ \cdot \\ \cdot \\ u^j(x) = \langle x, u^j \rangle = u^j_1x^1 + \ldots + u^j_ix^i + \ldots + u^j_nx^n = 0 \\ \cdot \\ \cdot \\ \cdot \\ u^p(x) = \langle x, u^j \rangle = u^p_1x^1 + \ldots + u_ix^i + \ldots + u^p_nx^n = 0. \end{cases}$$

Este sistema de relaciones es un *sistema de ecuaciones cartesianas* del subespacio $(F^*)^{\perp} \subset E$. Estudiaremos la recíproca de este resultado en el capítulo 10 (Ecuaciones lineales).

152. Transposición

a) Sea E y F dos espacios vectoriales sobre K, E^* y F^* sus duales respectivos, f una aplicación lineal de E en F, y^* una forma lineal definida sobre F, $y^* \circ f$ es una aplicación de E en K, es lineal (compuesta de dos

aplicaciones lineales, § 140, teorema 1); es, por tanto, una forma lineal x^* definida sobre E. Tenemos, pues, el diagrama siguiente

$$\begin{array}{ccc} E & \xrightarrow{f} & F \\ & {\scriptstyle x^*}\searrow \quad \swarrow{\scriptstyle y^*} & \\ & K & \end{array} \qquad x^* = y^* \circ f.$$

Por tanto, dada una aplicación lineal f de E en F, a toda forma lineal y^* definida sobre F se puede hacer corresponder una forma lineal única definida sobre E, o sea $x^* = y^* \circ f$, esta aplicación $y^* \to y^* \circ f$ de F^* en E^* se llama la transpuesta de f y se designa ${}^tf$ (se lee "transpuesta de f")

$$(\forall\, y^* \in F^*) \qquad {}^tf(y^*) = y^* \circ f$$

por tanto,

$$\boxed{(\forall\, x \in E) \qquad [{}^tf(y^*)]\,(x) = (y^* \circ f)\,(x) = y^*[f(x)]}$$

y, en consecuencia, ${}^tf$ está definida por

$$(\forall\, x \in E) \qquad (\forall\, y^* \in F^*) \qquad \langle x,\ {}^tf(y^*)\rangle = \langle f(x),\ y^*\rangle$$

Esta aplicación ${}^tf$ al aplicar el espacio vectorial F^* sobre el espacio vectorial E^* se puede preguntar si es lineal. Lo es, en efecto: sea y^*, z^* dos elementos cualesquiera de F^* y λ un elemento cualquiera de K, tendremos si se utilizan las propiedades de la forma bilineal canónica

$$\begin{aligned}[{}^tf(y^* + z^*)](x) &= \langle x,\ {}^tf(y^* + z^*)\rangle = \langle f(x),\ y^* + z^*\rangle \\ &= \langle f(x),\ y^*\rangle + \langle f(x),\ z^*\rangle = \langle x,\ {}^tf(y^*)\rangle + \langle x,\ {}^tf(z^*)\rangle \\ &= [{}^tf(y^*)]\,(x) + [{}^tf(z^*)]\,(x)\end{aligned}$$

$$\begin{aligned}[{}^tf(\lambda y^*)](x) &= \langle x,\ {}^tf(\lambda y^*)\rangle = \langle f(x),\ \lambda y^*\rangle = \lambda\langle f(x),\ y^*\rangle \\ &= \lambda\langle x,\ {}^tf(y^*)\rangle = \lambda[{}^tf(y^*)](x).\end{aligned}$$

Estas igualdades (en K) al tener lugar para todo x de E

$${}^tf(y^* + z^*) = {}^tf(y^*) + {}^tf(z^*), \qquad {}^tf(\lambda y^*) = \lambda\,{}^tf(y^*).$$

EJEMPLO

Sea $E = F$, $f = id_E$ tendremos

$$\langle x,\ {}^t(id_E)\,(y^*)\rangle = \langle id_E(x),\ y^*\rangle = \langle x,\ y^*\rangle$$

por tanto,

$${}^t(id_E)\,(y^*) = y^*,$$

es decir,

$${}^t(id_E) = id_{E^*}.$$

TEOREMA Y DEFINICIÓN. — *Si* E *y* F *son dos espacios vectoriales sobre el mismo cuerpo conmutativo* K *y* f *una aplicación lineal de* E *en* F, *la aplicación de* F^* *en* E^* *definida por* $y^* \to {}^t f(y^*) = y^* \circ f$ *es lineal, se la llama la* transpuesta *de* f*; está definida por*

$$(\forall x \in E) \qquad (\forall y^* \in F) \qquad \langle x, {}^t f(y^*) \rangle = \langle f(x), y^* \rangle.$$

Si E = F

$${}^t(\mathrm{id}_E) = \mathrm{id}_{E^*}.$$

b) **Propiedades de la transposición**

1. La transposición $f \to {}^t f$ hace, por tanto, corresponder a un elemento f de $\mathfrak{L}(E, F)$ un elemento ${}^t f$ de $\mathfrak{L}(F^*, E^*)$; vamos a ver que esta aplicación es *lineal*; es suficiente para ello aplicar la relación fundamental de definición. Sean f y g dos aplicaciones lineales de E en F, ${}^t f$ y ${}^t g$ sus transpuestas respectivas

$$\begin{aligned}\langle x, {}^t(f+g)(y^*) \rangle &= \langle (f+g)(x), y^* \rangle = \langle f(x) + g(x), y^* \rangle \\ &= \langle f(x), y^* \rangle + \langle g(x), y^* \rangle = \langle x, {}^t f(y^*) \rangle + \langle x, {}^t g(y^*) \rangle \\ &= \langle x, {}^t f(y^*) + {}^t g(y^*) \rangle = \langle x, ({}^t f + {}^t g)(y^*) \rangle\end{aligned}$$

$$\begin{aligned}\langle x, {}^t(\lambda f)(y^*) \rangle &= \langle (\lambda f)(x), y^* \rangle = \langle \lambda f(x), y^* \rangle = \lambda \langle f(x), y^* \rangle \\ &= \lambda \langle x, {}^t f(y^*) \rangle = \langle x, \lambda {}^t f(y^*) \rangle\end{aligned}$$

y como estas dos igualdades tienen lugar para todo x de E y todo y^* de F, tenemos para todo f, todo g de $\mathfrak{L}(E, F)$ y todo λ de K

$${}^t(f+g) = {}^t f + {}^t g, \qquad {}^t(\lambda f) = \lambda {}^t f.$$

2. Consideremos ahora tres espacios vectoriales E, F, G sobre el mismo cuerpo K, sus duales respectivos E^*, F^*, G^*, las aplicaciones lineales $f\colon E \to F$ y $g\colon F \to G$ y sus transpuestas ${}^t f$, ${}^t g$ tenemos los diagramas



Vamos a demostrar que ${}^t(g \circ f) = {}^t f \circ {}^t g$; aplicando siempre mecánicamente la fórmula fundamental a $g \circ f$, si x es un elemento cualquiera de E y z^* un elemento cualquiera de G^* (puesto que $g \circ f$ es una aplicación de E en G)

$$\alpha = \langle x, {}^t(g \circ f)(z^*) \rangle = \langle (g \circ f)(x), z^* \rangle = \langle g[f(x)], z^* \rangle$$

apliquemos la fórmula fundamental a g y ${}^t g$

$$\alpha = \langle g[f(x)], z^* \rangle = \langle f(x), {}^t g(z^*) \rangle$$

y finalmente a f y ${}^t f$

$$\alpha = \langle f(x), {}^t g(z^*) \rangle = \langle x, {}^t f[{}^t g(z^*)] \rangle = \langle x, ({}^t f \circ {}^t g)(z^*) \rangle$$

de donde

$$^t(g \circ f) = {}^tf \circ {}^tg.$$

Apliquemos esto a un isomorfismo f de E sobre F, f^{-1} es un isomorfismo de F sobre E y se tiene

$$f^{-1} \circ f = \mathrm{id}_E \qquad f \circ f^{-1} = \mathrm{id}_F$$

como hemos visto que $^t(\mathrm{id}_E) = \mathrm{id}_{E^*}$ tendremos, por tanto,

$$^tf \circ {}^t(f^{-1}) = \mathrm{id}_{E^*} \qquad {}^t(f^{-1}) \circ {}^tf = \mathrm{id}_{F^*}$$

en consecuencia, ^tf es inversible y $^t(f^{-1}) = (^tf)^{-1}$.

3. Busquemos finalmente la transpuesta de ^tf, o sea, $^t(^tf)$, tenemos

$$f \in \mathcal{L}(E,\ F) \qquad {}^tf \in \mathcal{L}(F^*,\ E^*), \qquad {}^t({}^tf) \in \mathcal{L}(E^{**},\ F^{**}).$$

Nos colocamos únicamente en el caso en que E *y* F *son de dimensión finita sobre* K, E^{**} y F^{**} canónicamente isomorfos, respectivamente, a E y a F están identificados a E y F; vamos a demostrar que $^t(^tf) = f$.

Si se aplica la fórmula fundamental a ^tf, tendremos cualquiera que sean y^* de F^* y x^{**} de E^{**}

$$\langle y^*,\ {}^t({}^tf)(x^{**}) \rangle = \langle ({}^tf)(y^*),\ x^{**} \rangle$$

de donde como consecuencia de la identificación y si se aplica la fórmula fundamental a ^tf

$$\langle y^*,\ {}^t({}^tf)(x) \rangle = \langle {}^tf(y^*),\ x \rangle = \langle y^*,\ f(x) \rangle$$

de donde $^t(^tf) = f$: así, en el caso en que E y F son de dimensión finita toda aplicación lineal de F^* en E^* se puede considerar como la transpuesta de una aplicación lineal de E en F; dicho de otro modo la aplicación $f \to {}^tf$ de $\mathcal{L}(E,\ F)$ en $\mathcal{L}(F^*,\ E^*)$ es suprayectiva (y, por tanto, biyectiva), pues $\dim \mathcal{L}(E,\ F) = \dim \mathcal{L}(F^*,\ E^*) = \dim E \dim F$ (ver § 144). Podemos recapitular estos resultados en el teorema siguiente:

Teorema 1. — *Si* $f \in \mathcal{L}(E,\ F)$ *y* $^tf \in \mathcal{L}(F^*,\ E^*)$ *la aplicación* $f \to {}^tf$ *es lineal, es decir,*

$$^t(f + g) = {}^tf + {}^tg, \qquad {}^t(\lambda f) = \lambda f.$$

2. *Si* $f \in \mathcal{L}(E,\ F)$, $\quad g \in \mathcal{L}(F,\ G)$, $\quad {}^tf \in \mathcal{L}(F^*,\ E^*)$, $\quad {}^tg \in \mathcal{L}(G^*,\ F^*)$

$$^t(g \circ f) = {}^tf \circ {}^tg.$$

Si f *es biyectiva, igualmente lo es* ^tf *y*

$$^t(f^{-1}) = (^tf)^{-1}.$$

3. *Si* E *y* F *son de dimensión finita,* $f \to {}^tf$ *es un isomorfismo y, además,*

$$^t(^tf) = f.$$

EJERCICIO

Si generalizamos la noción de forma bilineal, diremos que $(x_1, x_2) \to F(x_1, x_2)$ es una aplicación bilineal de $E_1 \times E_2$ en E_3 (E_1, E_2, E_3 espacios vectoriales sobre el mismo cuerpo K) si las aplicaciones parciales $F(., x_2)$ y $F(x_1, .)$ son lineales (§ 149, *a*), y § 164). Demostrar que la aplicación $(f, y^*) \to y^* \circ f$ es una aplicación bilineal de $\mathcal{L}(E, F) \times F^*$ en E^* y que ${}^tf$ no es otra que la aplicación parcial relativa a f.

153. Transposición y ortogonalidad

a) Sea f: $E \to F$ una aplicación lineal ${}^tf$: $F^* \to E^*$ su transpuesta, consideremos sus núcleos y sus imágenes

$$N = \operatorname{Ker} f \subset E \qquad I = \operatorname{Im} f \subset F$$
$$N' = \operatorname{Ker} {}^tf \subset F^* \qquad I' = \operatorname{Im} {}^tf \subset E^*$$

busquemos el ortogonal $I^\perp$ de I, es una parte de F^* descrita por y^*, tal que para todo x de E $\langle f(x), y^*)\rangle = 0$, de donde

$$\langle f(x), y^* \rangle = \langle x, {}^tf(y^*)\rangle = 0$$

es decir, ${}^tf(y^*)(x) = 0$, para todo x de E; por tanto, $I^\perp$ está descrito por los y^* tales que ${}^tf(y^*) = 0$, es decir, $I^\perp = N'$:

TEOREMA. — *Sea f una aplicación lineal de* E *en* F, *el ortogonal (en* F^*) *de la imagen de f es el núcleo de la transpuesta ${}^tf$ de f.*

b) Caso en que E y F son de dimensión finita sobre K

Pongamos $\dim E = \dim E^* = n$, $\dim F = \dim F^* = p$. Hay reciprocidad entre f y ${}^tf$; por tanto,

$$(\operatorname{Im} f)^\perp = \operatorname{Ker} {}^tf \qquad (\operatorname{Im} {}^tf)^\perp = \operatorname{Ker} f;$$

por otra parte, si se utilizan los resultados de los §§ 143 (teorema 7) y 151 *b*)

$$\operatorname{rg}({}^tf) = \dim F^* - \dim \operatorname{Ker} {}^tf = \dim F^* - \dim (\operatorname{Im} f)^\perp$$
$$= \dim F^* - (\dim F^* - \dim \operatorname{Im} f) = \dim \operatorname{Im} f = \operatorname{rg}(f).$$

TEOREMA. — *Si* E *y* F *son dos espacios vectoriales de dimensión finita sobre el mismo cuerpo conmutativo* K *y f una aplicación lineal de* E *en* F

$$\operatorname{rg}({}^tf) = \operatorname{rg}(f).$$

Ejercicios

141. Determinar, según los valores de α (o de α, β, γ) los rangos de los sistemas S_1, S_2 S_3 de vectores de $\mathbf{R}^3$

S_1: $a = (\alpha, 1, 1)$, $b = (1, \alpha, 1)$, $c = (1, 1, \alpha)$
S_2: $a = (\alpha, 1, 1)$, $b = (-1, -\alpha, -1)$, $c = (-1, -1, \alpha)$ (M.G.P.)
S_3: $a = (0, \gamma, -\beta)$, $c = (-\gamma, 0, \alpha)$, $c = (\beta, -\alpha, 0)$.

142. Determinar los rangos de los sistemas siguientes $\mathbf{R}^4$ y eventualmente las relaciones entre los vectores de S_i $(i = 1, 2, 3)$

S_1: $a = (1, 1, 1, 1)$, $b = (0, 1, 2, -1)$, $c = (1, 0, -2, 3)$, $d = (2, 1, 0, -1)$
S_2: $a = (1, 0, 1, 0)$, $b = (2, 1, 0, 1)$, $c = (0, 2, -1, 1)$, $d = (3, -1, 2, 0)$
S_3: $a = (1, 0, 2, 3)$, $b = (7, 4, -2, -1)$, $c = (5, 2, 4, 7)$, $d = (3, 2, 0, 1)$.

143. En $\mathbf{R}^n$ demostrar que el subespacio engendrado por a y b, por una parte, y el subespacio engendrado por c y d, de otras partes, son idénticas, y determinar su dimensión

1. $n = 3$ $a = (2,3 - 1)$, $b = (1, -1, -2)$, $c = (3, 7, 0)$, $d = (5, 0, -7)$
2. $n = 4$ $a = (2,3 - 1, 0)$, $b = (-3, 1, 0, 2)$, $c = (-4, 5, -1, 4)$, $d = (9, 8, -3, -2)$.

En cada uno de los casos anteriores, completar $\{a, b\}$ para obtener una base de $\mathbf{R}^3$ (caso 1) o de $\mathbf{R}^4$ (caso 2).

144. Determinar una base del subespacio V de $\mathbf{R}^4$ descrito por (x_1, x_2, x_3, x_4) del que $x_1 = x_2 - 3x_3$, $x_3 = x_4$. Completar la base obtenida para obtener una base de $\mathbf{R}^4$.

145. 1. En $\mathbf{R}^3$ verificar que a, b, c son independientes y calcular las coordenadas de x sobre la base $\{a, b, c\}$

$a = (-1, 1, 1)$, $b = (1, -1, 1)$, $c = (1, 1, -1)$, $x = (2, -3, -1)$.

2. El mismo problema en $\mathbf{C}^3$:

$a = (1, -1, i)$, $b = (-1, i, 1)$, $c = (i, 1, -1)$, $x = (1 + i, 1 - i, i)$.
$a = (1, 2, -1, -2)$, $b = (2, 3, 0, -1)$, $c = (1, 2, 1, 4)$, $d = (1, 3, -1, 0)$,
$x = (7, 14, -1, 2)$ (M.G.P).

146. Determinar los rangos de los sistemas de vectores siguientes de $\mathbf{R}^n$ (se estudiará primero el caso $n = 2$, 3 y 4).

1. $1 \leq i \leq n$), a_i tiene por coordenadas $\alpha_i^j = 1 + (\alpha - 1)\,\delta_i^j$
(α escalar dado, δ_i^j símbolo de Kronecker).
2. $1 \leq i \leq n$, a_i tiene por coordenadas $\alpha_i^j = (j - 1)\,n + i\,\alpha$ (α escalar dado).

147. En $\mathbf{R}^4$ se considera el subespacio F engendrado por $\{a, b, c\}$ y el subespacio G engendrado por $\{d, e\}$:

$a=(1, 2, 3, 4)$, $b=(2, 2, 2, 6)$, $c=(0, 2, 4, 4)$, $d=(1, 0, -1, 2)$, $e=(2, 3, 0, 1)$

Determinar las dimensiones de F, G, $F \cap G$, $F + G$ y las bases de estos subespacios (se buscará primero una base de $F \cap G$).

148. En $E = \mathcal{F}(\mathbf{R}, \mathbf{R})$ sobre $\mathbf{R}$, se considera f_k definida por $f_k(t) = e^{r_k t}$ $(r_k \in \mathbf{R})$. Demostrar que la familia (f_k) $(1 \leq k \leq n)$ es libre si y solamente si los r_k son todos distintos (razonar por inducción: escribir $\lambda_1 f_1(t) + \ldots + \lambda_n f_n(t) = 0$, dividir por $f_n(t)$ y derivar).

Aplicar este resultado a $\mathcal{F}(\mathbf{R}, \mathbf{C})$ con $r_k \in \mathbf{C}$.
(Se admitirá $r \in \mathbf{C}$, $t \in \mathbf{R}$: $(e^{rt})' = re^{rt}$).

149. Si p y q son dos números complejos $(q \neq 0)$, se considera el conjunto S de las sucesiones (u_n) $(n \in \mathbf{N})$ de los números tales que

$$u_{n+2} + pu_{n+1} + qu_n = 0 \qquad (n \in \mathbf{N})$$

se pone $(r \in \mathbf{C})$

$$(u_n) + (u'_n) = (u_n + u'_n) \qquad r(u_n) = (ru_n).$$

a) Demostrar que S es un subespacio vectorial del espacio vectorial, sobre $\mathbf{C}$, de todas las sucesiones complejas.

b) Se considera la aplicación de S en $\mathbf{C}^2$ que, a toda sucesión (u_n) de S, hace corresponder el elemento (u_0, u_1) de $\mathbf{C}^2$. Demostrar que esta aplicación es un isomorfismo de S sobre $\mathbf{C}^2$.

c) Demostrar que existe una sucesión única (a_n) de S tal que $a_0 = 1$, $a_1 = 0$ y una sucesión única (b_n) de S tal que $b_0 = 0$, $b_1 = 1$. De lo anterior deducir que para toda sucesión (u_n) de S existe un par único de números complejos (α, β) tal que $u_n = \alpha\, a_n + \beta\, b_n$ para todo n. ¿Cuál es la dimensión de S sobre $\mathbf{C}$?

d) Demostrar que si $p^2 - 4q \neq 0$, hay en S dos sucesiones linealmente independientes de la forma s^n $(s \in \mathbf{C})$. Calcular u_n en función de n, u_0, u_1, p, q.

Aplicación numérica. 1. $p = q = -1$, $u_0 = u_1 = 1$;
2. $p = -2k \cos\varphi$, $q = k^2$, $u_0 = 1$, $u_1 = k \cos\varphi$ $(k, \varphi$ reales).

e) Si $p^2 - 4q = 0$ sólo hay en S una sucesión de la forma s^n $(s \in \mathbf{C})$. ¿Qué relación verifica la sucesión (v_n) definida por $u_n = s^n v_n$ para todo n? Deducir de este estudio dos sucesiones de S linealmente independientes. Calcular u_n en función de n, u_0, u_1, p.
(M. G. P.).

150*. Se llamará sucesión (a_i) toda sucesión finita estrictamente creciente de números reales verificando

(1) $$0 = a_0 < a_i < \ldots < a_i < a_{i+1} < \ldots < a_n < a_{n+1} = 1$$

(no es necesario que n sea el mismo para cada sucesión).

Se designará por E el conjunto de las funciones reales en *escalera* definidas sobre $[0, 1]$, es decir las funciones f tales que existe un entero n, una sucesión (a_i) que verifica (1) y dos sucesiones (b_i) de $n + 1$ elementos reales cualesquiera tales que

(2) $$(i = 0, 1, \ldots, n) \qquad a_i \leq x < a_{i+1}, \qquad f(x) = b_i.$$

Se designa por L el conjunto de las funciones reales *lineales a «trozos»* definidas sobre $[0, 1]$, es decir, las funciones g tales que existe un entero n, una sucesión (a_i) que verifica (1) y dos sucesiones (b_i) y (c_i) cada una de $n + 1$ reales cualesquiera tales que

(3) $$(i = 0, 1, \ldots, n) \qquad a_i \leq x < a_{i+1}, \qquad f(x) = b_i x + c_i.$$

1.° a) Demostrar que E es un espacio vectorial sobre $\mathbf{R}$.

b) Dado el número real k $(0 \leq k < 1)$ se designa por e_k la función definida sobre [0, 1[por

$$0 \leq x < k, \qquad e_k(x) = 0; \qquad k \leq x < 1, \qquad e_k(x) = 1.$$

Demostrar que si (k_i) es una sucesión finita de números reales distintos dos a dos, la familia (e_{k_i}) es una familia libre de E.

c) Demostrar que toda función en escalera es una combinación lineal finita de funciones e_k y esto de una manera única.

2.° a) Demostrar que L es un espacio vectorial sobre **R** y que L′, conjunto de las funciones g lineales por trozos y *continuas* sobre [0, 1[y *nulas para cero* es un subespacio vectorial de L.

b) Dado el número real k $(0 \leq k < 1)$ se designa por l_k la función definida sobre [0, 1[por

$$0 \leq x < k, \qquad l_k(x) = 0; \qquad k \leq x < 1, \qquad l_k(x) = x - k.$$

Demostrar que si (k_i) es una sucesión finita de números reales distintos dos a dos, la familia (l_{k_i}) es una familia libre de L.

c) Demostrar que todo elemento de L′ es una combinación lineal finita de funciones (l_k) y esto de una manera única.

d) Demostrar que $L = E \oplus L'$.

3.° A toda función f de E se asocia la función h definida por

$$0 \leq x < 1, \qquad h(x) = \int_0^x f(t)\, dt.$$

a) Demostrar que $h \in L'$.

b) Se pone $h = \varphi(f)$, demostrar que φ es un isomorfismo de los espacios vectoriales sobre **R**, E y L′.

151. Se dice que un número complejo x es algebraico si verifica al menos una relación

(1) $$a_0 + a_1 x + \dots + a_n x^n = 0$$

(a^k coeficientes *racionales* no todos nulos).

a) Demostrar que, para que x sea racional, es necesario y suficiente que el subespacio vectorial V(x) de **C**, espacio vectorial sobre **Q**, engendrado por $\{1, x, \dots, x^n, \dots\}$ sea de dimensión finita. Si p es la dimensión de V(x) demostrar que existe una y una sola relación con coeficientes racionales de la forma

$$a_0 + a_1 x + \dots + a_{p-1} x^{p-1} + x^p = 0$$

Se dice que x es *algebraico de grado p*.

Demostrar que $2 - i\sqrt{2}$, $1 + \sqrt[3]{2}$ son algebraicos ¿cuál es su grado respectivo?

b) Si x es algebraico de grado p e y algebraico de grado q, demostrar que xy y $x + y$ son algebraicos. (Considerar el subespacio de **C**, espacio vectorial sobre **Q**, engendrado por $(x^k y^h)$ $(0 \leq h < p,\ 0 \leq k < q,\ h$ y k enteros naturales).

Demostrar que $\sqrt{2} + \sqrt[3]{3}$ es algebraico.

c) Demostrar que el conjunto de los numeros algebraicos es un subcuerpo del cuerpo **C** (M.G.P.).

152. Si E y F son dos espacios vectoriales sobre el mismo cuerpo conmutativo K, y f una aplicación de E en F, demostrar que la aplicación φ de $E \times F$ en sí mismo definida por $\varphi(x, y) = (x, y - f(x))$ es un automorfismo de $E \times F$.

153. Si E y F son dos espacios vectoriales sobre el cuerpo conmutativo K, de dimensión finita, expresados, respectivamente, en la base $\{a_1, ..., a_m\}$ y en la base $\{b_1, ..., b_n\}$, se consideran las mn aplicaciones f_j^i de E en F $(1 \leq i \leq m, 1 \leq j \leq n)$ definidas por

$$(k = 1, 2, ..., m) \quad f_j^i(a_k) = \delta_k^i b_j.$$

(δ_k^i símbolo de KRONECKER). Demostrar que los f_j^i son independientes y que engendran el espacio vectorial sobre K, $\mathfrak{L}(E; F)$. Deducir de lo anterior

$$\dim \mathfrak{L}(E; F) = \dim E \times \dim F.$$

(Se definirá f aplicación de E en F mediante $f(a_i) = \sum_{j=1}^{n} \alpha_i^j b_j$ y se demostrará que $f = \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_i^j f_j^i$)

154*. Si E y F son dos espacios vectoriales sobre el cuerpo conmutativo **K**, se supone que

$$E = E_1 \oplus ... \oplus E_i ... \oplus E_m, \qquad F = F_1 \oplus ... \oplus F_j \oplus ... \oplus F_n$$
$$x = x_1 + ... + x_i + ... + x_m, \qquad y = y_1 + ... + y_j + ... + y_n$$
$$(x \in E, \quad i = 1, ..., m, \quad x_i \in E_i; \qquad y \in F, \quad j = 1, ..., n, \quad y_j \in F_j).$$

A todo elemento f de $\mathfrak{L}(E; F)$ se hace corresponder la familia (f_i) $(1 \leq i \leq m)$: donde f_i es la restricción de f a E_i, la familia (f^j) $(1 \leq j \leq n)$, con $f^j = pr^j \circ f$ y finalmente la familia (f_i^j) con $f_i^j = pr^j \circ f_i$ (pr^j está definido por $pr^j(y) = y_j$).

a) Demostrar que f^j es una aplicación lineal de E en F_j tal que $f(x) = \sum_{j=1}^{n} f^j(x)$ para todo x. Recíprocamente si f^j es una aplicación lineal cualquiera de E en F_j, demostrar que g definido por $g(x) = \sum_{j=1}^{n} f^j(x)$ para todo x, es la única aplicación lineal de E en F tal que $pr^j \circ g = f^j$.
Demostrar que $\mathfrak{L}(E, F)$ es suma directa de los subespacios $\mathfrak{L}(E; F_j)$.
b) Demostrar que f_i es una aplicación lineal de E_i en F tal que $f = \sum_{i=1}^{n} f_i \circ pr^i$ (Si pr^i está definida por $pr^i(x) = x_i$). Recíprocamente si f_i es una aplicación lineal cualquiera de E_i en F $(1 \leq i \leq m)$, demostrar que existe una aplicación lineal única de E en F cuya restricción a cada E_i sea f_i (v. ej. 10 cap. 1).

Demostrar que $\mathfrak{L}(E; F)$ es isomorfo al espacio vectorial producto $\prod_{i=1}^{m} \mathfrak{L}(E_i; F)$.

c) Demostrar finalmente que f_i^j es una aplicación lineal de E_i en F_j tal que $f(x) = \sum_{i=1}^{m}\sum_{j=1}^{n} (f_i^j \circ pr^i)(x)$ para todo x

Recíprocamente demostrar que mn aplicaciones lineales cualesquiera f_i^j de E_i en F_j determinan así una y sólo una aplicación lineal f de E en F.

d) Sea G un tercer espacio vectorial sobre **K**, tal que

$$G = G_1 \oplus \dots \oplus G_k \oplus \dots \oplus G_p.$$

se pone $h = g \circ f$ y se designa por g_j^k (resp. h_i^k) la aplicación de F_j en G_k (resp. de E_i en G_k) asociada a g (resp a h) como f_i^j está asociado a f; demostrar que

$$h_i^k = \sum_{j=1}^{n} g_j^k \circ f_i^j.$$

155*. Si E es un espacio vectorial sobre el cuerpo conmutativo **K**, se designa por f un endomorfismo de E que permuta con *todo* automorfismo u de E.

a) Demostrar que, si f es $\neq 0$ y x no pertenece a Ker f, x e $y = f(x)$ son colineales (designando por F el subespacio engendrado por y, y por G un suplementario de F respecto a E, se demostrará que si x e y son independientes, u definido por $u(y) = x + y$, y $u(z) = z$ para todo z de G es un automorfismo de E tal que $f \circ u \neq u \circ f$). Deducir que existe un escalar $\lambda(x)$ tal que $f(x) = \lambda(x)x$.

b) Demostrar que $\lambda(x)$ es independiente de x (Considerar $x \neq x_2$ y un autoformismo u de E tal que $u(x_1) = x_2$).

c) ¿Cuál es el centro del anillo $\mathfrak{L}(E)$?

156. Sea E un espacio vectorial de dimensión finita n sobre el cuerpo conmutativo **K** y $\{a_1, a_2, \dots, a_n\}$ una base de E. Demostrar que r elementos ($r \leq n$) de E, $x_1, x_2, \dots, x_r$ son independientes si y solamente si existe un automorfismo f de E tal que $f(a_k) = x_k$ para todo k de $[1, r]$.

157. Si E es un espacio vectorial sobre un cuerpo conmutativo **K**, se llama *proyector* todo endomorfismo de E tal que $p^2 = p \circ p = p$. Se designa por e la identidad de E.

a) Demostrar que p es un proyector si y solamente si $e - p$ lo es. ¿Qué relaciones hay entre las imágenes y los núcleos de p y de $e - p$?

b) Demostrar que si p es un proyector

$$E = \text{Im } p \oplus \text{Ker } p.$$

(*Esta relación no caracteriza los proyectores,* ver ej. 159.)

c) Demostrar que un proyector p permuta con todo endomorfismo f tal que *Im* p y Ker p sean estables por f.

d) *Si p_1 y p_2 son dos proyectores, ¿qué condición será necesaria y suficiente* para que $p_1 + p_2$ también lo sean?

158. Si f es un endomorfismo de E espacio vectorial sobre un cuerpo conmutativo **K** de característica nula, demostrar que $f^2 = e$ ($e = \text{id}_E$) si y solamente si $1/2(f + e)$ es un proyector.

159. Sea f un endomorfismo de E de dimensión finita sobre el cuerpo conmutativo **K**, demostrar que

(1) $$E = \text{Im } f \oplus \text{Ker } f \Leftrightarrow \text{Im } f = \text{Im } (f^2).$$

Dar un ejemplo de un endomorfismo f de $\mathbf{R}^2$, espacio vectorial sobre $\mathbf{R}$, que no es un proyector.

160. Si f es un endomorfismo no inyectivo de un espacio vectorial E de dimensión n sobre un cuerpo conmutativo $\mathbf{K}$ se pone,

$$f^0 = \mathrm{id}_E,\ f^k = f^{k-1} \circ f \ (k \geq 1) \quad \text{y} \quad \mathrm{Ker}\ (f^k) = N_k,\ \mathrm{Im}\ (f^k) = I_k.$$

1.° Demostrar que para todo entero natural k

$$N_k \subset N_{k+1}, \qquad I_k \supset I_{k+1}.$$

2.° Demostrar que hay un entero natural p tal que

$$k < p \quad N_k \neq N_{k+1}, \qquad k \geq p \quad N_k = N_{k+1}.$$

Demostrar que $p \leq n$ y que

$$k < p \quad I_k \neq I_{k+1}, \qquad k \geq p \quad I_k = I_{k+1},$$

3.° Demostrar que I_p y N_p son dos subespacios suplementarios de E y que la restricción de f a I_p tomando sus valores en I_p es un automorfismo de I_p. (M.G.P.).

161. Sea f un endomorfismo de E de dimensión finita sobre un cuerpo K de característica $\neq 2$. Se propone demostrar que f es la diferencia de dos automorfismos de E; se demostrará esta propiedad sucesivamente en los casos siguientes:

a) f es un automorfismo de E (se observará que $f + f$ es también un automorfismo de E).

b) f es un endomorfismo no nulo y no inyectivo tal que $E = \mathrm{Im}\, f \oplus \mathrm{Ker}\, f$ (utilizar el ejercicio 159).

c) f es un endomorfismo cualquiera de E (se demostrará que existe un automorfismo a de E tal que $g = f \circ a$ verifica $E = \mathrm{Im}\ g \oplus \mathrm{Ker}\ g$).

162. Sea f un endomorfismo de un espacio vectorial de dimensión n sobre un cuerpo conmutativo; demostrar que Ker f = Im f si y solamente si $f^2 = 0$, $f \neq 0$, n es par y rg $f = n/2$.

163. Si $E_0, E_1, \ldots, E_n$ son espacios vectoriales sobre el mismo cuerpo conmutativo K ($n \geq 2$) se dice que el diagrama

$$E_0 \xrightarrow{f_0} E_1 \to \ldots \to E_{k-1} \xrightarrow{f_{k-1}} E_k \xrightarrow{f_k} E_{k+1} \to \ldots \to E_{n-1} \xrightarrow{f_{n-1}} E_n$$

es una *sucesión exacta* si para $0 \leq k \leq n-2$

$$\mathrm{Im}\ f_k = \mathrm{Ker}\ f_{k+1}.$$

Si E_0 (resp. E_n) es igual a $\{0\}$ que se escribirá O, no se escribe f_0 (resp. f_{n-1}) pues hay una sola aplicación lineal de O en E_1 (resp. de E_{n-1} en O).

a) Demostrar

$$[O \to E \xrightarrow{f} F \text{ es una sucesión exacta}] \Leftrightarrow f \text{ es inyectiva.}$$

$$[E \xrightarrow{f} F \to O \text{ es una sucesión exacta}] \Leftrightarrow f \text{ es suprayectiva.}$$

b) Demostrar que los diagramas siguientes son sucesiones exactas

$$O \to F \xrightarrow{i} E \xrightarrow{s} E/F \to O$$

$$O \to \operatorname{Ker} f \xrightarrow{i} E \xrightarrow{f} F \xrightarrow{s} F/\operatorname{Im} f \to O$$

(f aplicación lineal, i inyección canónica, s homorfismo suprayectivo canónico).

164. a) Dados $n+1$ espacios vectoriales de dimensión finita sobre el mismo cuerpo conmutativo K, se considera la sucesión exacta (ver ej. 163)

$$O \to E_0 \xrightarrow{f_0} \ldots \to E_k \xrightarrow{f_k} E_{k+1} \to \ldots \xrightarrow{f_{n-1}} E_n \to O.$$

Demostrar que

$$\sum_{2h+1 \leq n} \dim E_{2h+1} = \sum_{2h \leq n} \dim E_{2h}, \qquad \sum_{k=0}^{n} (-1)^k \dim E_k = 0.$$

b) Demostrar que el diagrama siguiente es una sucesión exacta (E y F espacios vectoriales de dimensión finita)

$$O \to E \cap F \xrightarrow{f+g} E \oplus F \xrightarrow{i-j} E + F \to O$$

donde $f : E \cap F \to E$ y $g : E \cap F \to F$ son inyecciones canónicas, igualmente que $i : E \to E + F$ y $j : F \to E + F$.
Deducir de lo anterior que (V. § 143, ej. 1)

$$\dim (E \cap F) + \dim (E + F) = \dim E + \dim F.$$

165. Sea E una álgebra de dimensión n sobre el cuerpo conmutativo K.

a) Si $\{a_1, \ldots, a_n\}$ es una base de E, considerado como espacio vectorial sobre **K**, demostrar que existe para (i, j) describiendo $[1, n] \times [1, n]$ los escalares λ_{ij}^k **tales** que

$$(1) \qquad a_i a_j = \sum_{k=1}^{n} \lambda_{ij}^k a_k$$

verificando

$$(2) \qquad \sum_{l=1}^{n} \lambda_{ij}^l \lambda_{lk}^m = \sum_{l=1}^{n} \lambda_{il}^m \lambda_{jk}^l$$

(utilizar $(a_i a_j) a_k = a_i (a_j a_k)$). Las fórmulas (1) definen la *«tabla de multiplicar»* del álgebra E.

b) Recíprocamente dado un espacio vectorial E sobre el cuerpo conmutativo **K** de base $\{a_1, \ldots, a_n\}$ y una familia de escalares λ_{ij}^k que verifica las relaciones (2), demostrar que las fórmulas

$$x = \sum_{i=1}^{n} \alpha^i a_i, \qquad y = \sum_{j=1}^{n} \beta^j a_j, \qquad xy = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \alpha^i \beta^j \lambda_{ij}^k a_k$$

dan a E una estructura de álgebra sobre K (se verificará primero la distributividad del producto respecto a la suma y se deducirá la asociatividad de la multiplicación de la de la multiplicación de los elementos de la base).

c) ¿Qué propiedades verifica la tabla de multiplicar de E si un elemento de la base, a_1 por ejemplo, es igual a un elemento unidad e de **K**? (Se supone $\mathbf{K} \subset E$ por identificación, ver § 147, ej. 7.)

d) Demostrar que toda extensión cuadrática $\mathbf{K}[\alpha]$ de un cuerpo conmutativo **K** (ver cap. 5, ej. 96 y 97) es un álgebra sobre K. ¿Cuál es su dimensión?

166. Se designa por D el conjunto de pares de *números reales* $A = (a, a'), \ldots,$ $X = (x, x'), \ldots,$ provistos de las dos operaciones:

(1) $$A + B = (a, a') + (b, b') = (a + b, a' + b').$$
(2) $$AB = (a, a')(b, b') = (ab, ab' + a'b).$$

Todo elemento de D se llama *número dual*. (¡No tiene ninguna relación con el dual de un espacio vectorial!)

a) Demostrar que D′ descrito por $(x, 0)$ es estable para las operaciones (1) y (2) y que D′ provisto de las leyes inducidas es isomorfo a **R**: se identificará D′ y **R** poniendo $(x, 0) = x$.

b) Demostrar que D es una álgebra sobre **R** de dimensión 2 y que $1 = (1, 0)$ y $\alpha = (0, 1)$ forman una base de D. Calcular α^2.

c) Se pone $X = (x, x') = x + \alpha x'$, $\overline{X} = x - \alpha x'$, $f(X) = \sqrt{X\overline{X}}$; ¿la aplicación f de D en $\mathbf{R}_+$ es una norma?

d) Calcular $(x + \alpha x')^n$ (n entero, $n \geq 2$). Discutir y resolver las ecuaciones

$$X^2 = A, \qquad X^2 - 2PX + Q = 0, \qquad X^n = A.$$

(A, P, Q números duales dados, X dual incógnita). (M. G. P.)

167. En el espacio vectorial $\mathbf{R}^4$ sobre **R** expresado en su base canónica: $e = (1, 0, 0, 0)$, $i = (0, 1, 0, 0)$, $j = (0, 0, 1, 0)$, $k = (0, 0, 0, 1)$ se define una multiplicación poniendo

$$e^2 = e,\ i^2 = j^2 = k^2 = -e,\ ei = ie = i,\ ej = je = j,\ ek = ke = k$$
$$jk = -kj = i,\ ki = -ik = j,\ ij = -ji = k.$$

a) Demostrar que se define así una algebra de dimensión 4 sobre **R** (ver. ej. 165): se llama el álgebra K de los *cuaterniones*. Demostrar que los elementos $(s, 0, 0, 0)$ describen una subálgebra de K isomorfa a **R**: se la identificará con **R** poniendo $e = (1, 0, 0, 0) = 1$.

b) Demostrar que todo cuaternión $x = (s, a, b, c)$ se escribe de una manera única: $x = s + ai + bj + ck$. Se pone

$$xx' = (s + ai + bj + ck)(s' + a'i + b'j + c'k) = S + Ai + Bj + Ck.$$

Calcular S, A, B, C en función de $s, a, b, \ldots, c'$.

Se escribe $x = (s, \vec{v})$, $x' = (s', \vec{v'})$, $xx' = (S, \vec{V})$ (s, s', S reales, $\vec{v}$, $\vec{v'}$, $\vec{V}$ vectores de $\mathbf{R}^3$), calcular S y V (se utilizará el producto escalar y el producto vectorial de v y v').

Ponemos $\bar{x} = s - ai - bj - ck$ (si $x = s + ai + bj + ck$) y $N(x) = x\bar{x}$. Demostrar que $N(x)$ es real y positivo y que $N(x) = 0$ equivale a $x = 0$. Demostrar que $N(xy) = N(x)N(y)$ (se observará que $x\bar{x} = \bar{x}x$). De lo anterior deducir que **K** es

un cuerpo (no conmutativo). Distinguir el subgrupo de **K*** engendrado por *i, j, k* (V. capítulo 4, ej. 78).

d) Todo cuaternión $x \neq 0$ puede escribirse $x = r(\cos\theta, \vec{u} \operatorname{sen}\theta)$ si $\vec{u}$ es un vector de $\mathbf{R}^3$ tal que $\|\vec{u}\| = 1$, y r real estrictamente positivo. ¿Es única esta representación?

Calcular $x^n = \rho(\cos\varphi, \vec{v} \operatorname{sen}\varphi)$ $(\rho > 0, \|\vec{v}\| = 1)$.
Resolver las ecuaciones con incógnita x en **K** $x^2 = -1$, $x^2 = i$. (M. G. P.)

168 Si E es un álgebra sobre el cuerpo conmutativo K, se llama *derivación* en E todo endomorfismo D de E, considerado como espacio vectorial sobre **K**, tal que

$$(\forall x, y \in \mathrm{E}) \qquad \mathrm{D}(xy) = \mathrm{D}(x)y + x\mathrm{D}(y).$$

Se designará el conjunto de las derivaciones de E por $\mathfrak{D}(\mathrm{E})$. e es el elemento unidad de **K**, se supone por identificación: $\mathrm{K} \subset \mathrm{E}$ (§ 147, ej. 7). Finalmente se escribe $\mathrm{D}^0 = \mathrm{id}_\mathrm{E}$, $\mathrm{D}^n = \mathrm{D}^{n-1} \circ \mathrm{D}$.

a) Calcular $\mathrm{D}(e)$, $\mathrm{D}(\alpha)$ $(\alpha \in \mathbf{K})$ y $\mathrm{D}(y)$ con

$$y = \alpha_0 + \alpha_1 x \ldots + \alpha_n x^n \quad (\alpha_0, \ldots, \alpha_n \in \mathrm{K},\ x \in \mathrm{E}).$$

Calcular $\mathrm{D}^p(y)$. ¿Qué obtendremos si **K** es de característica no nula?

b) Demostrar que (C_n^k coeficiente binomial)

$$\mathrm{D}^n(xy) = \sum_{k=0}^{n} \mathrm{C}_n^k \mathrm{D}^k(x)\, \mathrm{D}^{n-k}(y).$$

c) Demostrar que si $a \in \mathrm{E}$, $x \mapsto ax - xa$ es una derivación en E. (Observar que en E la multiplicación no es forzosamente conmutativa, ver ej. 167.)

d) Demostrar que $\mathfrak{D}(\mathrm{E})$ es un subespacio vectorial del espacio vectorial sobre **K**, $\mathcal{L}(\mathrm{E})$.

e) Demostrar que si D_1 y D_2 son derivaciones en E, $\mathrm{D}_1 \circ \mathrm{D}_2$ no lo es, pero que $[\mathrm{D}_1, \mathrm{D}_2] = \mathrm{D}_1 \circ \mathrm{D}_2 - \mathrm{D}_2 \circ \mathrm{D}_1$ si que lo es.

169. Sea E un espacio vectorial sobre el cuerpo conmutativo **K**. Demostrar que p formas lineales $f^1 \ldots f^p$ definidas sobre E son independientes si y solamente si existe x de E tal que para todo i de $[1, p]$, $f(x) = \alpha^i$, siendo $\alpha^1 \ldots \alpha^p$ escalares cualesquiera (considerar la aplicación g de E en $\mathbf{K}^p$ definida por

$$g(x) = (f^1(x), \ldots, f^p(x)).$$

170. Estudiar la independencia de las formas lineales definidas sobre $\mathbf{R}^4$ y que tienen como valores

1. $x_1 - \lambda x_3$, $x_2 - \dfrac{1}{\lambda} x_4$, $x_1 - \mu x_4$, $x_2 - \dfrac{1}{\mu} x_4$ $\quad (\lambda, \mu \in \mathbf{R}^*)$

2. $x_1 - x_3 \operatorname{sen}\alpha - x_4 \cos\alpha$, $x_2 - x_3 \cos\alpha + x_4 \operatorname{sen}\alpha$,
$x_1 + x_3 \operatorname{sen}\beta - x_4 \cos\beta$, $x_2 - x_3 \cos\beta - x_4 \operatorname{sen}\beta$ $\quad (\alpha, \beta \in \mathbf{R})$.

OBSERVACION

Todos los ejercicios desde el 141 hasta el 147 pueden considerarse como ejercicios sobre las formas lineales definidas sobre $\mathbf{K}^n$ considerando las n-étuplas dadas como si definiesen las coordenadas de formas lineales respecto a la base de $(\mathbf{K}^n)^*$ dual de la base canónica de $\mathbf{K}^n$.

171*. Se propone demostrar que si E es un espacio vectorial de dimensión $n > 1$ sobre un cuerpo conmutativo **K**, no existe ningún isomorfismo canónico φ de E sobre E*, salvo si $n = 2$ y **K** es isomorfismo a $\mathbf{Z}/2\mathbf{Z}$. (Se recuerda que un isomorfismo canónico sólo debe depender de la estructura de espacio vectorial y de ningún modo de la elección de la base.)

a) Demostrar que si φ existe para todo automorfismo u de E y todo x y todo y de E, se debe tener

$$(1) \qquad \langle x, \varphi(y) \rangle = \langle u(x), (\varphi \circ u)(y) \rangle.$$

b) Si $n > 1$ y $\mathbf{K} \neq \mathbf{Z}/2\mathbf{Z}$, demostrar que existe un automorfismo u de E no verificando (1). (Tomar u tal que $u(x) = \lambda x$, $\lambda \neq 0$, $\lambda \neq 1$, $u(y) = y$.)

c) Si $n > 2$ y **K** cualquiera, demostrar que existe un automorfismo u de E que no verifica (1). (Si y no es nulo y designando por F el subespacio de E ortogonal a $\varphi(y)$, demostrar que se puede tomar $x \notin F$, $u(x) \in F$, $u(y) = y$.)

d) Estudiar el caso del espacio vectorial $\mathbf{K}^2$ sobre el cuerpo $\mathbf{K} = \mathbf{Z}/2\mathbf{Z}$.

172*. Sea E un espacio vectorial sobre el cuerpo conmutativo **K** admitiendo una base (a_i) $(i \in \mathrm{I})$ infinita y E* su dual. Para todo x de E, $x = \sum_{i \in \mathrm{I}} \alpha^i a_i$ sólo un número finito de escalares α_i son no nulos. Se designa siempre por a^{*i} la forma lineal tal que $a^{*i}(x) = \alpha^i$.

a) Demostrar que (a^{*i}) $(i \in \mathrm{I})$ es una familia libre de E* pero que no engendra E*. (Considerar la forma lineal f tal que para todo i de I, $f(a_i) = 1$.)

b) Demostrar que el homomorfismo $x \to \tilde{x}$ de E en E** es inyectivo y no suprayectivo. Deducir que la parte del bidual descrito por $\tilde{x}$, cuando x describe E, es un subespacio propio del bidual.

173. Si E es un espacio vectorial de dimensión n sobre el cuerpo conmutativo **K**, referido en una base (a_i), E* su dual referido en la base dual (a^{*i}), se considera el isomorfismo φ de E sobre E* definido por $\varphi(a_i) = a^{*i}$ para todo $i \in [1, n]$. Hallar todos los automorfismos u de E tales que, para todo x y todo y de E, se tenga

$$\langle x, \varphi(y) \rangle = \langle u(x), (\varphi \circ u)(y) \rangle$$

(se definirá u mediante los escalares u_i^j tales que

$$u(a_i) = \sum_{j=1}^{n} u_i^j a_j \quad (1 \leq i \leq n,\ 1 \leq j \leq n)$$

y se buscará las relaciones entre las u_i^j).

174. Sea F y G dos subespacios vectoriales de un espacio vectorial E de dimensión finita sobre el cuerpo conmutativo **K**, demostrar que

$$(F + G)^{\perp} = F^{\perp} \cap G^{\perp}, \qquad (F \cap G)^{\perp} = F^{\perp} + G^{\perp}.$$

De lo anterior deducir que: $E = F \oplus G \Rightarrow E^* = F^{\perp} \oplus G^{\perp}$.

175. En $E = \mathbf{R}^4$ espacio vectorial sobre **R**, se considera el subespacio F engendrado por (1, 1, 1, 1), (—1, 1, —2, 2), (—1, 5, —4, 8), (—3, 1, —5, 3).

a) ¿Cuál es la dimensión de F? ¿Cuál es la dimensión en E^* de $F' = F^{\perp}$? Demostrar que la imagen de $v = (x, y, z, t)$ de E para todo forma lineal $f \in F'$ puede escribirse: $f(v) = 4ax + 4by - (3a + b)z - (a + 3b)t$. De lo anterior deducir dos formas f_1 y f_2 que constituyan una base de E^* dual de la base canónica de E.

b) Demostrar que la relación «$f = g$ sobre F» es una relación de equivalencia $\mathcal{R}$ definida sobre E^* y que es compatible con la estructura del espacio vectorial de E^*. Demostrar que $\mathcal{R}$ es equivalente a «$f - g \in F'$».

Demostrar que la imagen de una base de F por un elemento φ del espacio cociente E^*/F' determina φ. Deducir de ello que E^*/F' es isomorfo al dual F^* de F.

176. Si E y F son dos espacios vectoriales sobre el mismo cuerpo conmutativo **K**, si V es un subespacio vectorial de E y f una aplicación lineal de E en F, demostrar que

$$[f(V)]^{\perp} = ({}^{t}f^{-1}(V^{\perp})).$$

177. Si E y F son dos espacios vectoriales de dimensión finita sobre el mismo cuerpo conmutativo **K**, si V es un subespacio vectorial de E, se designa por $\mathcal{L}_V(E; F)$ el conjunto de las aplicaciones lineales de E en F que se anulan sobre V. Se designará por W un suplementario cualquiera de V.

a) Demostrar que $\mathcal{L}_V(E; F)$ es un subespacio vectorial de $\mathcal{L}(E; F)$ isomorfo a $\mathcal{L}(E/V; F)$ y a $\mathcal{L}(W; F)$.

b) Demostrar que $E^* = V^{\perp} \oplus W^{\perp}$ (ver ej. 174) y que V^* y W^* son, respectivamente, isomorfos a $W^{\perp}$ y $V^{\perp}$.

c) Demostrar que si f es una aplicación lineal de E en F, ${}^{t}f(F^*)$ es isomorfo a $[f(E)]^*$.

178. Si E y F son dos espacios vectoriales de igual dimensión finita sobre el cuerpo conmutativo **K** y f un isomorfismo de E sobre F, se pone

$$\check{f} = ({}^{t}f)^{-1} = {}^{t}(f^{-1})$$

que se lee «*f tchèche*». (V. § 152, **b** 2).

a) Demostrar que $\check{f}$ es un isomorfismo de E^* sobre F^* correspondiendo canónicamente a f.

b) Si g es también un isomorfismo de E sobre un tercer espacio vectorial G

$$\left(\check{\overline{g \circ f}}\right) = \check{g} \circ \check{f}.$$

Deducir que si f es un automorfismo de E (de dimensión finita), la aplicación $f \to \check{f}$ define un isomorfismo de GL(E) sobre GL(E^*).

179. Si en la definición de un espacio vectorial sobre un cuerpo conmutativo **K** se sustituye **K** por un *anillo conmutativo unitario* A, se obtiene la definición de un A-*módulo*.

Las *definiciones* de las nociones siguientes: módulo-producto, submódulo, módulo-cociente, parte generatriz, parte libre, parte ligada, bases, homomorfismos, etc., se obtienen trasladando las definiciones relativas a los espacios vectoriales. Pero las *propiedades* de los módulos pueden ser *muy diferentes* de las propiedades correspondientes de los espacios vectoriales: sólo subsisten íntegramente las que, de hecho, sólo utilizan la estructura de anillo del cuerpo **K**. Por el contrario *un módulo no tiene forzosamente una base* y todo submódulo de un módulo M no tiene forzosamente suplementario en M.

Se llama módulo *libre* todo módulo que posee al menos una base y módulo de *tipo finito* todo módulo que admite una parte generatriz finita, de donde tenemos cuatro categorías de módulos: 1. módulo libre de tipo finito; 2. módulo libre; 3. módulo de tipo finito; 4. módulo cualquiera.

a) En un A-módulo se puede tener $\lambda x = 0$, $\lambda \neq 0$, $x \neq 0$; el resultado del teorema 1 (§ 133) puede ser falso; un elemento $x \neq 0$ no es forzosamente una parte libre. (Observar que A es un A-módulo).

b) Demostrar que $\mathbf{Z}^n$ es un **Z**-módulo libre de tipo finito (admite la base canónica $e_i = (\delta_i^1, \ldots, \delta_i^i, \ldots, \delta_i^n)$, $i = 1, \ldots, n$).

c) Demostrar que todo grupo abeliano representado aditivamente es un **Z**-módulo. Demostrar que todo grupo abeliano finito M (representado aditivamente) es un **Z**-módulo de tipo finito no libre ($n =$ card. M, si hubiera una base $\{a_1, \ldots, a_m\}$, $(m \leq n$, M sería isomorfo a $\mathbf{Z}^m$ que es infinito).

d) Demostrar que los submódulos de **Z**, considerado como **Z**-módulo, son los conjuntos $p\mathbf{Z}$ ($p \in \mathbf{N}$). Demostrar que el submódulo $p\mathbf{Z}$ ($p \neq 0$) no tiene suplementario en **Z**.

e) Tomando de nuevo los teoremas de este capítulo relativos a los espacios vectoriales sobre un cuerpo conmutativo, indicar los que se pueden trasladar a cada una de las cuatro categorías de módulos indicados anteriormente. Por ejemplo: $\mathcal{L}(M, N)$ conjunto de los homomorfismos de dos A-módulos (A anillo conmutativo unitario) es un A-módulo.

180. La teoría de los espacios vectoriales sobre **K** (resp. los A-módulos, A anillo unitario) puede aplicarse al caso en que **K** (resp. A) no es conmutativo.

Un espacio vectorial sobre **K** (resp. un A-módulo) en el que las dos operaciones verifican los ocho axiomas del § 125 es un *espacio vectorial por la izquierda* (resp. un A-*módulo por la izquierda*). Si la operación externa es $(\lambda, x) \to x\lambda$ con $xe = x$, $(x\mu)\lambda = x(\mu\lambda)$, $x(\lambda + \mu) = x\lambda + x\mu$, $(x + y)\lambda = x\lambda + y\lambda$, se dice que se tiene un *espacio vectorial por la derecha* (resp. un A-*módulo por la derecha*).

a) Si E y F son dos espacios vectoriales por la izquierda sobre **K** (resp. dos A-módulos por la izquierda), demostrar que $\mathcal{L}(E; F)$ es un espacio vectorial por la izquierda sobre el centro C de **K** (resp. un C-módulo por la izquierda, si C es el centro del anillo A).

b) Si E es un espacio vectorial por la izquierda sobre **K** (resp. un A-módulo por la izquierda), el dual de E, $E^* = \mathcal{L}(C; K)$ (resp. $E^* = \mathcal{L}(E; A)$ es un espacio vectorial por la derecha sobre **K** (resp. un A-módulo por la derecha).

8

MATRICES

I. Generalidades

154. Definiciones diversas

a) DEFINICIÓN. — *Dados dos conjuntos finitos de índices* I *y* J *y un conjunto* E *se llama* matriz *de tipo* (I, J) *sobre* E *toda aplicación de* I × J *en* E.

Una matriz es, pues, una *familia doble finita* (§ 17)

$$(i, j) \to a_{ij} \qquad \text{o} \qquad (i, j) \to a_i^j$$

ordinariamente se colocan los elementos a_{ij} (o a_i^j) según una tabla rectangular: uno de los índices es el índice de la columna del cuadro, el otro el de la fila. Por razones que explicaremos en el § 155, *escogeremos en general en este curso la notación* a_i^j, *i es el índice de la columna y j el índice de la fila.* Una matriz se representará ya por (a_i^j) $((i, j) \in I \times J)$, ya por el *cuadro* de los elementos de a_i^j, cuadro puesto entre dos paréntesis (ciertos autores ponen corchetes) o dos rayas verticales en la izquierda y derecha del cuadro, también si no da lugar a confusión por una sola letra A.

Dos matrices A y B definidas, respectivamente, por

$$\begin{array}{llll} A: & I \times J \to E & & (i, j) \to a_i^j \\ B: & I' \times J' \to E & & (i, j) \to b_i^j \end{array}$$

serán iguales si y sólo si

$$I = I', \qquad J = J', \qquad (\forall (i, j) \in I \times J) \qquad a_i^j = b_i^j$$

Cada vez que no lo precisemos tomaremos

$$I = [1, m], \qquad J = [1, n], \qquad m = \text{card } I, \qquad n = \text{card } I.$$

$$A = (a_i^j)_{\substack{(1 \le i \le m)\\(1 \le j \le n)}} = \begin{pmatrix} a_1^1 & a_2^1 & \dots & a_i^1 & \dots & a_m^1 \\ a_1^2 & a_2^2 & \dots & a_i^2 & \dots & a_m^2 \\ \vdots & \vdots & & \vdots & & \vdots \\ a_1^j & a_2^j & \dots & a_i^j & \dots & a_m^j \\ \vdots & \vdots & & \vdots & & \vdots \\ a_1^n & a_2^n & \dots & a_i^n & \dots & a_m^n \end{pmatrix}$$

Diremos que A es una matriz de *tipo* (m, n) sobre E (*m número de columnas, n de filas*), $M_E(m, n)$ o $M(m, n)$ si no hay lugar a confusión.

Si $m = n$, se dice que la matriz es *cuadrada de orden n.*
Si $m = 1$, se dice que la matriz es *unicolumna.*
Si $n = 1$, se dice que la matriz es *unifila.*

Sea A la muestra (a_i^j) $((i, j) \in I \times J)$, si $I' \subset I$ y $J' \subset J$ la restricción a $I' \times J'$ de la aplicación $(i, j) \to a_i^j$ es una *submatriz* A′ de A: **se** obtiene conservando en A únicamente las columnas cuyo índice pertenece a **I′** y las filas cuyo índice pertenece a J′, o suprimiendo en A las columnas cuyo índice pertenece a $\complement_I I'$ y las filas cuyo índice pertenece a $\complement_J J'$. Para A′, en general es $I' \neq [1, m']$ si $m' = \text{card } I'$, sino que $I' = \{i_1, i_2, \dots, i_m\}$ con

$$1 \le i_1 < i_2 \dots < i_{m'} \le m$$

e igualmente para J′.

Inversamente cuando se pasa de A′ a A se dice que se *orla* la matriz A′ por las columnas de índice i pertenecientes a $\complement_I I'$ y las líneas de índices j pertenecientes a $\complement_J J'$.

Si $I' = [a, b]$, $J' = [c, d]$ $(1 \le a < b \le m,\ 1 \le c < d \le n)$ se dice que la submatriz A′ es un *bloc.*

Se llama *transpuesta* de la matriz $A = (a_i^j)$ $((i, j) \in I \times J)$, la matriz $B = (b_i^j)$ $((i, j) \in J \times I)$ definida por

$$b_i^j = a_j^i$$

se la representa[32] ${}^tA$ que se lee "*transpuesta de* A". Si A es de tipo (m, n), ${}^tA$ es de tipo (m, n), por ejemplo,

$${}^t\begin{pmatrix} a & b & c \\ a' & b' & c' \end{pmatrix} = \begin{pmatrix} a & a' \\ b & b' \\ c & c' \end{pmatrix}$$

$${}^t(x_1, x_2, \dots, x_n) = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

(32) Según los autores se usan otras notaciones: A′, $\widetilde{A}$, ..., para la transpuesta de A.

La transposición $A \to {}^tA$ es, pues, una aplicación de $M(m, n)$ en $M(n, m)$, es evidentemente biyectiva; además,

$${}^t({}^tA) = A.$$

Se dice que una matriz es *simétrica* si $A = {}^tA$, esto implica que $m = n$, luego A es cuadrada y

$$(\forall i \in I)(\forall j \in I) \qquad a^i_j = a^j_i.$$

Los elementos a^i_i de una matriz cuadrada se llaman *elementos diagonales* de A, su conjunto es la *diagonal principal* de A; el conjunto de los elementos a^j_i $(i + j = n + 1)$ se llama algunas veces la diagonal no principal de A. Decir que una matriz es simétrica quiere decir que dos elementos colocados simétricamente en relación a la diagonal principal son iguales.

b) Las matrices se utilizan en las más diversas ramas de las actividades humanas, para el matemático son útiles principalmente si se puede efectuar con ellas cálculos por medio de sus elementos. Con el propósito de simplificar supondremos en lo sucesivo (salvo mención contraria) que E es un *cuerpo conmutativo,* los elementos neutros se representan 0 y 1.

Sin necesidad de repetirlo $M(m, n)$ designará en adelante el conjunto de las matrices de m columnas y n filas sobre un cuerpo conmutativo K (que a menudo en las aplicaciones será **R** o **C**); en lugar de $M_K(n, n)$, se escribe $M_n(K)$.

c) Para mayor comodidad agrupamos a continuación ciertas definiciones cuya justificación aparecerá en los párrafos siguientes:

Una matriz de tipo (m, n) es *nula* si y sólo si todos sus elementos son nulos. Se la representará O si no cabe ninguna confusión: pero la matriz O de tipo (m, n) no es igual a la O de tipo (m', n') más que si se verifica $m' = m$, $n' = n$.

Si $A = (a^i_j)$ $(1 \leq i \leq m,\ 1 \leq j \leq n)$ la matriz $(-a^j_i)$ del mismo tipo que A, se llama la *opuesta* de A y se representa $-A$.

Una matriz tal que ${}^tA = -A$ se llama *antisimétrica*: es cuadrada (A es de tipo (m, n), ${}^tA$ es de tipo (n, m)) y

$$(\forall i \in I)(\forall j \in I) \qquad a^i_j = -a^j_i$$

en particular

$$a^i_i = -a^i_i \Rightarrow 2a^i_i = 0$$

es decir, $a^i_i = 0$ si el cuerpo K no es de característica 2 (§ 97, 104).

Si $K = \mathbf{C}$ (resp. **R**) se dice que la matriz $A = (a^j_i)$ $(1 \leq i \leq m,\ 1 \leq j \leq n)$ es *compleja* (resp. *real*). Si A es compleja, la matriz $(\bar{a}^j_i)$ se llama la *conjugada* de A y se representa $\bar{A}$. Si $\bar{A} = A$ la matriz A es real.

En lo sucesivo A es la matriz cuadrada (a^j_i) de *orden n*, salvo indicación contraria. Se dice que A es *triangular inferior* si

$$i > j \Rightarrow a^j_i = 0$$

y *triangular superior* si

$$i < j \Rightarrow a^j_i = 0$$

A se llama matriz *diagonal* si $i \neq j$ $a^i_j = 0$ (una matriz diagonal es a la vez triangular inferior y triangular superior).

A se llama *matriz escalar* si $a^i_j = \delta^i_j a$, donde (δ^i_j) es el símbolo de KRONECKER (§ 133, ej. 1), es decir, es una matriz diagonal en la que todos los elementos de la diagonal principal son iguales entre sí. En particular la matriz (δ^i_j) de orden n se llama, veremos por qué, *matriz unidad* de orden n y se representa I_n.

Estas matrices, triangular inferior, diagonal, unidad, se escriben, respectivamente,

$$\begin{pmatrix} a^1_1 & & & & \text{O} \\ a^2_1 & a^2_2 & & & \\ \vdots & & & & \\ a^i_1 & a^i_2 & \dots & a^i_i & \\ \vdots & \vdots & & & \\ a^n_1 & a^n_2 & \dots & a^n_i & \dots\ a^n_n \end{pmatrix} \begin{pmatrix} a^1_1 & & & \text{O} \\ & a^2_2 & & \\ & & \ddots & \\ & & a^i_i & \\ \text{O} & & & \ddots \\ & & & a^n_n \end{pmatrix} \qquad I_n = \begin{pmatrix} 1 & & & \\ & \ddots & & \text{O} \\ & & 1 & \\ \text{O} & & & \ddots \\ & & & 1 \end{pmatrix}$$

el signo "O" no designa en las matrices anteriores un "bloc", pero simbolizando el hecho que todos los elementos estrictamente por encima —o debajo de la diagonal principal— son nulos; al contrario si se escribe, cuando A es de tipo (m, n),

$$A = \begin{pmatrix} I_r & \text{O} \\ \text{O} & \text{O} \end{pmatrix}$$

I_r y los tres "O" representan los blocs; el O debajo de I_r es una matriz nula de tipo $(r, n-r)$ y las otras dos (de arriba abajo) de tipo $(m-r, r)$, $(m-r, n-r)$.

Siendo A una *matriz compleja* (a^i_j), rectangular o cuadrada se llama *adjunta* de A, que se representa A^*, la matriz ${}^t(\bar{A}) = (\overline{{}^tA})$. Si A es tal que $A^* = A$ se dice que A es una matriz *hermitiana*, entonces es cuadrada; se tiene

$$a^i_j = \overline{a^j_i}$$

luego los elementos de la diagonal principal son reales $(a^i_i = \overline{a^i_i})$ y los elementos simétricos en relación a la diagonal principal son conjugados: se dice también que A presenta la simetría hermitiana.

EJEMPLOS

1. Siendo E un espacio vectorial de dimensión n sobre K de base $a_1, a_2, \dots, a_n$ y p vectores $x_1, x_2, \dots, x_p$

$$x_i = \sum_{j=1}^{n} \alpha^j_i a_j \qquad (i = 1, 2, \dots, p)$$

las coordenadas α_i^j de estos p vectores pueden estar dispuestas según una matriz

$$\begin{pmatrix} \alpha_1^1 \dots \alpha_i^1 \dots \alpha_p^1 \\ \vdots \quad \vdots \quad \vdots \\ \alpha_1^j \dots \alpha_i^j \dots \alpha_p^j \\ \vdots \quad \vdots \quad \vdots \\ \alpha_1^n \dots \alpha_i^n \dots \alpha_p^n \end{pmatrix}$$

es la *matriz de las coordenadas* de estos p vectores, la columna i correspondiendo al vector x_i. En particular la matriz unicolumna de las coordenadas α_i de x en una base (a_i) $(1 \leq i \leq n)$ se llamará la *matriz asociada a x relativamente a la base* (a_i); se la representa $M(x, (a_i))$ o $M(x)$ o también X, si no hay lugar a confusión.

2. Sobre este mismo espacio se puede también definir q formas lineales $u^1, u^2, \dots, u^q$ por las fórmulas (ver § 148)

$$x = \sum_{i=1}^{n} \alpha^i a_i, \qquad u^j(x) = \sum_{i=1}^{n} \lambda_i^j \alpha^i$$

poniendo $u^j(a_i) = \lambda_i^j$.

Las coordenadas de estas formas (en la base dual de la de E) pueden estar dispuestas según la matriz

$$\begin{pmatrix} \lambda_1^1 \dots \lambda_i^1 \dots \lambda_n^1 \\ \vdots \quad \vdots \quad \vdots \\ \lambda_1^j \dots \lambda_i^j \dots \lambda_n^j \\ \vdots \quad \vdots \quad \vdots \\ \lambda_1^q \dots \lambda_i^q \dots \lambda_n^q \end{pmatrix}$$

es la matriz de las coordenadas de estas q formas lineales, la línea j representando la forma u^j. En particular la matriz de una fila de las coordenadas λ^i de una forma lineal u en la base dual (a_i^*) $(1 \leq i \leq n)$ se llamará matriz asociada en la forma lineal u relativamente a la base (a_i^*).

3. De una manera general, dada una matriz A de tipo (m, n) sobre K, toda matriz unicolumna extraída de A se llama *vector-columna* de A (es un vector de K^n) y toda matriz de una fila de A se llama *vector-fila* de A (es un vector de K^m).

4. Dados dos espacios vectoriales en K de dimensiones respectivas m y n, de bases $\{a_1, \dots, a_m\}$ y $\{b_1, \dots, b_n\}$, se puede definir sobre $E \times F$ una *forma bilineal* f por los datos de m, n elementos α_{ij} de K definidos por

$$f(a_i, b_j) = \alpha_{ij}$$

estos m, n elementos de K pueden estar dispuestos según una matriz A de este tipo (m, n) (ver capítulo 15, § 221).

5. Finalmente, y es la interpretación más importante de una matriz sobre un cuerpo conmutativo K, se puede representar por una matriz toda *aplicación lineal* de E en F, E y F espacios vectoriales de dimensiones finitas sobre K, provistas de bases bien determinadas. El resto de este capítulo está dedicado a la relación entre matrices y aplicaciones lineales.

155. Matriz y aplicación lineal

a) Matriz asociada a una aplicación lineal

Sea E y F dos espacios vectoriales sobre K de dimensiones respectivas m y n, $\{a_1, a_2, \ldots, a_m\}$ una base de E y $\{b_1, b_2, \ldots, b_n\}$ una base de F y finalmente una aplicación lineal de E en F definida por

(1) $$x \longrightarrow y = f(x).$$

Si se pone(33)

$$x = \sum_{i=1}^{m} x^i a_i, \qquad y = \sum_{j=1}^{n} y^j b_j, \qquad f(a_i) = \sum_{j=1}^{n} \alpha_i^j b_j$$

tendremos, gracias a la linealidad de f y a las reglas de cálculo en un espacio vectorial sobre un cuerpo conmutativo

$$y = f(x) = \sum_i x^i f(a_i) = \sum_i \left[\sum_j x^i \alpha_i^j b_j \right]$$

$$= \sum_j b_j \left[\sum_i \alpha_i^j x^i \right] = \sum_j y^j b_j$$

de donde

(1′) $$y^j = \sum_{i=1}^{m} \alpha_i^j x^i \qquad (j = 1, 2, \ldots, n)$$

sistema que se puede escribir de manera desarrollada

(1″) $$\begin{cases} y^1 = \alpha_1^1 x^1 + \ldots + \alpha_i^1 x^i + \ldots + \alpha_m^1 x^m \\ \vdots \\ y^j = \alpha_1^j x^1 + \ldots + \alpha_i^j x^i + \ldots + \alpha_m^j x^m \\ \vdots \\ y^n = \alpha_1^n x^1 + \ldots + \alpha_i^n x^i + \ldots + \alpha_m^n x^m \end{cases}$$

en consecuencia, a toda aplicación lineal f de E en F, E expresado respecto a una base (a_i) $(1 \leq i \leq m)$ y F en una base (b_j) $(1 \leq j \leq n)$, podemos asociar una matriz $A = (\alpha_i^j)$ $(1 \leq i \leq m,\ 1 \leq j \leq n)$ por las fórmulas

(2) $$f(a_i) = \sum_{j=1}^{n} \alpha_i^j b_j \qquad (i = 1, 2, \ldots, m)$$

(33) No hemos querido abusar de las letras griegas (ver § 125), aquí x^i, y^j son escalares. Por otra parte, el lugar de los índices i y j, recuerda que i, índice inferior en α_i^j, es el número del vector $f(a_i)$ y que j, índice superior en α_i^j, es el número de las coordenadas de $f(a_i)$ en la base (b_j).

En fin, para simplificar escribimos algunas veces $\sum_i$ en lugar de $\sum_{i=1}^{m}$, si no da lugar a ninguna confusión.

dando las columnas de A sobre la base de F, las coordenadas de las imágenes por f de los vectores de la base de E.

Se escribe

$$A = M(f, (a_i), (b_j))$$

y si no da lugar a confusión, $A = M(f)$.

Fijadas las bases (a_i) y (b_j), hemos definido la aplicación M

$$M: \mathcal{L}_K(E, F) \to M_K(m, n).$$

Esta aplicación es biyectiva: en efecto (§ 143, *a*), existe una y sólo una aplicación f: $E \to F$ tal que $f(a_i) = c_i$, según esto toda matriz (α_i^j) de tipo (m, n) define, de una manera única, m vectores de F por las fórmulas

$$c_i = \sum_{j=1}^{n} \alpha_i^j b_j \qquad (i = 1, 2, \dots, m)$$

luego:

TEOREMA. — *Dada una aplicación f de un espacio vectorial E sobre K de base (a_i) $(1 \leq i \leq m)$, en un espacio vectorial F sobre K de base $(b_j)(1 \leq j \leq n)$ la aplicación*

$$f \to M((f, (a_i), (b_j)) = (\alpha_i^j)$$

definida por

$$f(a_i) = \sum_{j=1}^{n} \alpha_i^j b_j \qquad (i = 1, 2, \dots, m)$$

es una biyección de $\mathcal{L}_K(E, F)$ sobre $M_K(m, n)$, donde las columnas de M(f) tienen por elementos las coordenadas en la base (b_j) de F, de las imágenes por f de los vectores (a_i) de la base de E. M(f) se llama matriz asociada a f.

En particular, toda matriz (α_i^j) sobre K de tipo (m, n) puede ser considerada como asociada a la aplicación lineal de K^m en K^n referidos a sus bases canónicas con $x = (x^1, x^2, \dots, x^m)$, $y = (y^1, y^2, \dots, y^n)$ y

$$(j = 1, \dots, n) \qquad y^j = \sum_{i=1}^{m} \alpha_i^j x^i.$$

Observemos que cuando se utiliza una matriz para estudiar una aplicación lineal, no es necesario buscar la imagen de un vector cualquiera de E, sino únicamente las imágenes de los vectores de la base de E (ver teorema anterior y § 143, teorema 6).

Cuando E = F (f es, en consecuencia, un endomorfismo de E), la matriz asociada a f es cuadrada. En general se toma la misma base para E considerado como espacio de salida y espacio de llegada, se escribe entonces

$$A = M(f, (a_i), (a_i)) = M(f, (a_i))$$

pero se puede tomar también dos bases distintas. Veremos en la sección III (cambio de bases) los casos en que para id_E: $E \to E$ se emplea, por la misma naturaleza del problema, una base para E espacio de salida y otra para E espacio de llegada, naturalmente si se tomase la misma base

$$M(id_E, (a_i)) = I_n.$$

Esta biyección entre $\mathfrak{L}(E, F)$ y $M(m, n)$ (fijadas unas determinadas bases en E y F) es muy importante: permite pasar de toda noción o toda operación definida sobre las aplicaciones lineales a una noción u operación definida sobre las matrices: siendo los razonamientos y cálculos equivalentes en $\mathfrak{L}(E, F)$ y $M(m, n)$ (E y F de dimensiones finitas). Hay que notar, sin embargo:

1. En *matemáticas puras* es mejor razonar y calcular en $\mathfrak{L}(E, F)$; por un lado, los razonamientos y cálculos son intrínsecos (independientes de las bases escogidas); por otro lado, desde el punto de vista técnico, todo es mucho más simple, se evitan todas las complicaciones de escritura debidas a la superposición de índices. En fin, los razonamientos y cálculos en $\mathfrak{L}(E, F)$ son más generales: muchos de ellos se aplican a los espacios de dimensión infinita.
2. En *matemáticas aplicadas*, en los distintos grados del cálculo numérico es necesario servirse de las matrices.

Finalmente, surge una pregunta: ¿Cómo se transforma $M(f, (a_i), (b_j))$ cuando se cambian las bases? Estudiaremos este problema en la sección III.

b) Matrices asociadas a f y a ${}^tf$

Sea E y F espacios vectoriales sobre K de bases respectivas (a_i) $(1 \leq i \leq m)$ y (b_j) $(1 \leq j \leq n)$, sus duales, E^*, expresado en la base dual (a^{*i}) de la de E, F^*, expresado en la base dual (b^{*j}) de la de F (§ 150, *a*). Sea en fin f una aplicación lineal de E en F y ${}^tf$ su transpuesta, pongamos

$$A = M(f, (a_i), (b_j)), \qquad f(a_i) = \sum_{j=1}^{n} \alpha_i^j b_j \qquad (i = 1, 2, \ldots, m)$$

$$B = M({}^tf, (b^{*j}), (a^{*i})), \qquad {}^tf(b^{*j}) = \sum_{i=1}^{m} \beta_i^j a^{*i} \qquad (j = 1, 2, \ldots, n).$$

Observemos que en $B = (\beta_i^j)$, i es el índice de la *fila* y j el de la *columna*, esto para que en la última fórmula el índice de suma i (índice nulo) sea una vez índice inferior y otra vez índice superior. En $A = (\alpha_i^j)$, i es el índice de la *columna* y j el de la *fila*.

Dicho lo anterior, ${}^tf(b^{*j}) = b^{*j} \circ f$ (ver § 152) es un elemento de E^*, es decir, una forma lineal definida sobre E, que está definida por sus valores para a_i $(1 \leq i \leq m)$, de donde utilizando la fórmula fundamental de la dualidad

$${}^tf(b^{*j})(a_i) = \langle a_i, {}^tf(b^{*j}) \rangle = \langle f(a_i), b^{*j} \rangle = \left\langle \sum_{h=1}^{n} \alpha_i^h b_h, \; b^{*j} \right\rangle = \sum_{h=1}^{n} \alpha_i^h \delta_h^j = \alpha_i^j$$

pues $\langle b_h, b^{*j}\rangle = \delta^j_h$ (símbolo de KRONECKER) (ver § 150). (Se observará que i y j tienen valores fijos; debemos designar el índice de suma para toda letra, salvo i o j; por ejemplo, por h. Por otra parte,

$$^tf(b^{*j})(a_i) = \langle a_i, {}^tf(b^{*j})\rangle = \left\langle a_i, \sum_{k=1}^{m} \beta^j_k a^{*k} \right\rangle = \sum_{k=1}^{m} \beta^j_k \delta^k_i = \beta^j_i$$

de donde, *teniendo en cuenta la observación anterior sobre la significación de los índices en* α^j_i *y* β^j_i.

TEOREMA. — *Si* E *y* F *son dos espacios vectoriales de dimensiones finitas sobre* K, E *y* E* *están referidos a dos bases duales, así como* F *y* F*, *para todo elemento* f *de* $\mathfrak{L}(E, F)$ *se tiene*

$$M(^tf) = {}^t[M(f)].$$

EJEMPLOS Y EJERCICIOS

1. Sea f: $E \to F$, dim $E = m$, dim $F = n$, rg $f = r$; tomemos la notación del corolario 3 (§ 143): $a_1, a_2, \ldots, a_r, a_{r+1}, \ldots, a_m$ una base de E con $a_{r+1}, \ldots, a_m$ una base de Ker f, $b_1 = f(a_1), \ldots, b_r = f(a_r), b_{r+1}, \ldots, b_n$ es una base de F, se tiene entonces

$$M(f, (a_i), (b_j)) = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

matriz de cuatro bloques que hemos estudiado en el § 154, c.

2. En $\mathbf{R}^3$ expresado en una base $(\vec{i}, \vec{j}, \vec{k})$ la proyección sobre $\mathbf{R}^2$ (base $\vec{i}, \vec{j}$), paralelamente a $\vec{k}$, tiene por matriz asociada

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Con los mismos datos encontrar la matriz asociada a la proyección de $\mathbf{R}^3$ sobre $\mathbf{R}^2$ paralelamente al vector (p, q, r).

3. En $\mathbf{R}^2$ referido a una base ortonormal la rotación $\mathcal{R}(0, \theta)$ es asociada a la matriz (ver § 123, a)

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix}.$$

En $\mathbf{R}^3$ referido a una base ortonormal $(\vec{i}, \vec{j}, \vec{k})$ la rotación $\mathcal{R}(\vec{k}, \theta)$ está asociada a la matriz

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

4. En $\mathbf{R}^2$ referido a una base ortonormal $(\vec{i}, \vec{j})$ determinar la matriz asociada a la simetría en relación a la recta D tal que $(Ox, OD) \equiv \alpha \pmod{\pi}$.

5. En el espacio vectorial $\mathcal{F}_3$ los polinomios en x de coeficientes reales de grado ≤ 3, expresado en la base canónica $\{x^0, x^1, x^2, x^3\}$ (aquí 0, 1, 2, 3 son exponentes, $x^0 = 1$), encontrar la matriz asociada al endomorfismo $P \to P'$ (derivación) (ver § 1.., ej. 2, y capítulo 11, § 188).

II. Operaciones algebraicas de las matrices

Todas las matrices consideradas tienen, salvo indicación contraria, sus elementos en un cuerpo conmutativo K.

156. Espacio vectorial M(m, n)

a) Grupo aditivo M(m, n)

Sea $A = (\alpha_i^j)$, $B = (\beta_i^j)$ dos matrices de tipo (m, n) sobre K asociadas, respectivamente, a dos aplicaciones lineales f y g de E en F, E expresado en la base (a_i) $(1 \leq i \leq m)$ y F en la base (b_j) $(1 \leq j \leq n)$. Ponemos por definición

$$A + B = M(f, (a_i), (b_j)) + M(g, (a_i), (b_j)) = M(f + g, (a_i), (b_j)) = (\gamma_i^j)$$

la fórmula

$$(f + g)(a_i) = \sum_{j=1}^{n} \gamma_i^j b_i = f(a_i) + g(a_i) = \sum_{j=1}^{n} \alpha_i^j b_j + \sum_{j=1}^{n} \alpha_i^j b_j$$

demuestra que para todo i de $[1, m]$ y todo j de $[1, n]$ es

$$\gamma_i^j = \alpha_i^j + \beta_i^j.$$

Esta operación interna definida sobre M(m, n) se llama *adición* de matrices. Siendo las bases fijas la fórmula anterior, que se puede escribir

$$(1) \qquad M(f + g) = M(f) + M(g),$$

demuestra que la biyección $f \rightarrow M(f)$ es un isomorfismo del grupo aditivo $\mathcal{L}$(E, F) (§ 144) sobre el conjunto M(m, n) provisto de la adición de las matrices; este último es, pues, un grupo para la adición (§ 77, observación que sigue al teorema 2); en consecuencia,

$$O = M(o) \qquad -A = M(-f)$$

lo que justifica las nociones de *matriz nula* (de tipo (m, n)) y de *opuesta* de una matriz dada (§ 154, *c*).

b) Espacio vectorial M(m, n)

Para todo λ de K ponemos por definición

$$\lambda A = \lambda M(f, (a_i), (b_j)) = M(\lambda f, (a_i), (b_j)) = \lambda_i^j$$

la fórmula

$$(\lambda f)(a_i) = \sum_{j=1}^{n} \lambda_i^j b_j = \lambda f(a_i) = \lambda \sum_{j=1}^{n} \alpha_i^j b_j = \sum_{j=1}^{n} \lambda \alpha_i^j b_j$$

muestra que para todo i de $[1, m]$ y todo j de $[1, n]$

$$\lambda_i^j = \lambda\alpha_i^j.$$

Esta operación externa definida sobre $M(m, n)$, siendo el dominio de operadores K, se llama *multiplicación de matrices por un escalar.*

La fórmula (1) de más arriba y la fórmula que acabamos de obtener

$$M(\lambda f) = \lambda M(f) \tag{2}$$

demuestra que la biyección $f \to M(f)$ es un isomorfismo del espacio vectorial $\mathfrak{L}(E, F)$ sobre $M(m, n)$, luego:

TEOREMA. — *Si* E *y* F *son dos espacios vectoriales de dimensiones respectivas* m *y* n *sobre* K, $\mathfrak{L}(E, F)$, *provisto de las operaciones* $(f, g) \to f + g$, $(\lambda, f) \to \lambda f$ y $M(m, n)$, *provisto de las operaciones* $(A, B) \to A + B$ *y* $(\lambda, A) \to \lambda A$ *son espacios vectoriales isomorfos.*

Por otra parte, se ve fácilmente que

$$^t(A + B) = {}^tA + {}^tB$$
$$^t(\lambda A) = \lambda(^tA)$$

luego la aplicación $A \to {}^tA$ de $M(m, n)$ en $M(n, m)$ es un homomorfismo de espacios vectoriales; como es biyectivo (§ 154, *a*), es también un isomorfismo de espacios vectoriales. Vamos a ver, de otra manera, que los espacios vectoriales $M(m, n)$ y $M(n, m)$ son isomorfos.

c) **Base canónica y dimensión** $M(m, n)$

Sea la matriz $A = (\alpha_i^j)$ de tipo (m, n), tenemos

$$A = \sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i^j E_j^i$$

siendo E^i una matriz (m, n) en la que todos los elementos son nulos, salvo el elemento situado en la intersección de la columna i y de la fila j e igual a 1; en consecuencia,

$$E_j^i = \begin{pmatrix} & 0 & \\ O & \vdots & O \\ & 0 & \\ 0 \;\ldots\; 0 & 1 & 0 \;\ldots\; 0 \\ & 0 & \\ O & \vdots & O \\ & 0 & \end{pmatrix} \leftarrow \text{fila } j$$

$$\uparrow$$
columna i

Se sigue que estas m, n matrices E_j^i engendran $M(m, n)$; además, son visiblemente independientes; en efecto,

$$\sum_i \sum_j \alpha_i^j E_j^i = (\alpha_i^j)_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} = O \Rightarrow \alpha_i^j = 0$$

para todo i y todo j, luego:

TEOREMA Y DEFINICIÓN. — *El conjunto* (E_j^i) $(1 \leq i \leq m,\ 1 \leq j \leq n)$ *es una base de* $M(m, n)$, *llamada base canónica de* $M(m, n)$.

$M(n, m)$ *y* $M(m, n)$ *son dos espacios vectoriales isomorfos de dimensión* mn.

Encontramos así (ver ej. 153, fin del capítulo 7) que

$$\dim \mathcal{L}(E, F) = (\dim E)(\dim F).$$

El hecho de que $M(m, n)$ es un espacio vectorial permite decir que los α_i^j son las *coordenadas* de la matriz $A = (\alpha_i^j)$ (en relación a la base canónica (E_j^i)). Todas las nociones relativas a los elementos de un espacio vectorial pueden entonces aplicarse a las matrices de tipo (m, n): independencia lineal, subespacio vectorial, parte generatriz, etc., veremos ejemplos en los ejercicios siguientes.

EJERCICIOS

1. Sea S el conjunto de las matrices simétricas y A el conjunto de las matrices antisimétricas de $M_n(K)$ (el cuerpo K no es de característica 2), demostrar que S y A son dos subespacios suplementarios de $M_n(K)$. ¿Cuáles son sus dimensiones?

2. Sea $\mathcal{T}_s$ el conjunto de las matrices triangulares superiores de $M_n(K)$ y $\mathcal{T}_i$ el conjunto de las matrices triangulares inferiores. Demostrar que $\mathcal{T}_s$ y $\mathcal{T}_i$ son dos subespacios vectoriales isomorfos de $M_n(K)$, ¿cuál es su dimensión? Determinar $\mathcal{T}_s \cap \mathcal{T}_i$ y $\mathcal{T}_s + \mathcal{T}_i$.

3. Si a, b, c son tres números complejos cualesquiera, demostrar que las matrices

$$\begin{pmatrix} a+b & a-b+c & a-c \\ a-b-c & a & a+b+c \\ a+c & a+b-c & a-b \end{pmatrix}$$

describen un subespacio vectorial $M_3(C)$; indicar una base de este subespacio. ¿Cuál es su dimensión?

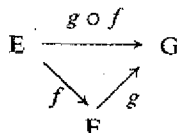
4. Si $t \to \alpha_i^j(t)$ son m, n funciones reales de la variable real t, derivables en $[t_1, t_2]$, demostrar que la matriz $A(t) = (\alpha_i^j(t))$, considerada como función vectorial (ver curso de Análisis) es derivable en $[t_1, t_2]$ y que

$$\frac{dA(t)}{dt} = \left(\frac{d\alpha_i^j(t)}{dt} \right).$$

157. Producto de matrices

a) Cálculo de la matriz producto de dos matrices

Sea tres espacios vectoriales E, F, G de dimensiones respectivas m, n, p sobre K, y de bases respectivas (a_i), (b_j), (c_k); consideremos las tres aplicaciones lineales f, g, $g \circ f$



y las tres matrices asociadas

$$A = M(f, (a_i), (b_j)) = (\alpha_i^j)$$
$$B = M(g, (b_j), (c_k)) = (\beta_j^k)$$
$$C = M(g \circ f, (a_i), (c_k)) = (\gamma_i^k)$$

ponemos por definición

$$C = BA = M(g, (b_j), (c_k))M(f, (a_i), (b_j)) = M(g \circ f, (a_i), (c_k))$$

observemos primero que

$$B = M(g) \in M(n, p), \qquad A = M(f) \in M(m, n), \qquad C = BA = M(g \circ f) \in M(m, p)$$

y que

$$\text{card }\{\text{columnas de B}\} = \text{card }\{\text{filas de A}\}$$

las fórmulas

$$(g \circ f)(a_i) = \sum_{k=1}^{p} \gamma_i^k c_k, \qquad f(a_i) = \sum_{j=1}^{n} \alpha_i^j b_j, \qquad g(b_j) = \sum_{k=1}^{p} \beta_j^k c_k$$

nos dan para todo i de $[1, m]$

$$(g \circ f)(a_i) = g[f(a_i)] = g\left[\sum_{j=1}^{n} \alpha_i^j b_j\right] = \sum_{j=1}^{n} \alpha_i^j g(b_j) = \sum_{j=1}^{n} \alpha_i^j \sum_{k=1}^{p} \beta_j^k c_k$$

$$= \sum\sum_{k} \alpha_i^j \beta_j^k c_k = \sum_{k}\sum_{j} \alpha_i^j \beta_j^k c_k = \sum_{k} c_k \sum_{j} \alpha_i^j \beta_j^k$$

utilizando las propiedades que traducen la estructura de espacio vectorial de G; de donde para todo i de $[1, m]$ y todo k de $[1, p]$

$$\gamma_i^k = \sum_{j=1}^{n} \beta^k \alpha_i^j$$

regla que se puede representar por el esquema siguiente

$$\underset{\longleftarrow m \longrightarrow}{p\updownarrow\begin{pmatrix} \\ \gamma_i^k \\ \\ \end{pmatrix}} = \underset{\longleftarrow n \longrightarrow}{p\updownarrow\begin{pmatrix} \\ \beta_1^k \cdots \beta_j^k \cdots \beta_n^k \\ \\ \end{pmatrix}} \underset{\longleftarrow m \longrightarrow}{\begin{pmatrix} \alpha_i^1 \\ \vdots \\ \alpha_i^k \\ \vdots \\ \alpha_i^n \end{pmatrix}} \updownarrow n$$

El elemento γ_i^k de BA (k número de la fila, i número de la columna) proviene de los elementos de la fila k de B (factor de la izquierda)[34] y de los elementos de la columna i de A (factor de la derecha)[34]; esta línea de B y esta columna de A tienen el mismo número de elementos: la dimensión del espacio intermediario F, que aquí es n. Además,

card {filas de BA} = card {filas de B}
card {columnas de BA} = card {columnas de A}.

Se dice que se ha efectuado el producto BA FIlas por COlumnas (en abreviatura "producto FICO").

A modo de ejercicio, busquemos el elemento general γ_i^j de $C = AB$, $A = (\alpha_i^j)$, $B = (\beta_i^j)$; para que este producto exista, según la regla anterior es necesario y suficiente que

card {elementos de una FIla de A} = card {elementos de una COlumna de B},

de donde

card {COlumnas de A} = card {FIlas de B},

es decir,

$$A \in M(n, m), \qquad B \in M(p, n),$$

se tendrá

$$\gamma_i^j = \sum_{k=1}^{n} \alpha_k^j \beta_i^k$$

y

$$C = AB \in M(p, m).$$

Apliquemos, siempre como "gimnasia" sobre los índices, la regla precedente con otras convenciones para el lugar de los índices:

1.° $A = (\alpha_i^j)$, $B = (\beta_i^j)$, $AB = (\gamma_i^j)$ i designa ahora el número de la fila y j el de la columna

$$\gamma_i^j = \sum_{k=1}^{n} \alpha_i^k \beta_k^j .$$

(34) B escrito en primer lugar (de izquierda a derecha, naturalmente) es asociada a la aplicación g efectuada en segundo lugar (ver observación de los §§ 15 y 44).

2.° $A = (\alpha_{ij})$, $B = (\beta_{ij})$, $AB = (\gamma_{ij})$ siendo i el primer índice el número de la fila y j el de la columna

$$\gamma_{ij} = \sum_{k=1}^{n} \alpha_{ik}\beta_{kj}$$

n es siempre en los dos ejemplos precedentes el número de las columnas de A y el de las filas de B.

b) **Propiedades de la multiplicación de matrices**

Las propiedades de la composición de las aplicaciones lineales y la biyección $f \to A = M(f)$ (§ 155) nos dan las propiedades correspondientes de la multiplicación de las matrices:

1. Sea el diagrama

$$E \xrightarrow{f} F \xrightarrow{g} G \xrightarrow{h} H$$

con $\dim_K E = m$, $\dim_K F = n$, $\dim_K G = p$, $\dim_K H = q$; escogidas unas bases en E, F, G, H pongamos

$$M(f) = A, \qquad M(g) = B, \qquad M(h) = C$$

tendremos

$$(h \circ g) \circ f = h \circ (g \circ f) \Rightarrow (CB)A = C(BA)$$

se escribirá

$$(CB)A = CBA \quad (\text{ver § } 45).$$

2. Por otra parte, si f_1 y f_2 son elementos de $\mathfrak{L}(E, F)$ y g_1 y g_2 elementos de $\mathfrak{L}(F, G)$, A_1, A_2 y B_1, B_2 las matrices asociadas, tendremos

$$(g_1 + g_2) \circ f \Rightarrow (B_1 + B_2)A = B_1A + B_2A$$
$$g \circ (f_1 + f_2) \Rightarrow B(A_1 + A_2) = BA_1 + BA_2.$$

Observemos que en general, es sólo por abuso de lenguaje que se puede llamar asociatividad y distributividad a estas propiedades, pues la aplicación $(B, A) \to BA$ no es, en general, una ley de composición interna: es una aplicación de $M(m, n) \times M(n, p)$ en $M(m, p)$ (ver § 145).

3. Busquemos si AB y BA pueden estar definidas simultáneamente; haría falta que se tuviera a la vez

$$\text{card}\,\{\text{columnas de A}\} = \text{card}\,\{\text{filas de B}\}$$
$$\text{card}\,\{\text{columnas de B}\} = \text{card}\,\{\text{filas de A}\},$$

luego $A \in M(m, n)$ y $B \in M(n, m)$ en este caso AB y BA existen

$$AB \in M(n, n), \qquad BA \in M(m, m)$$

si $m \neq n$, AB y BA no son, pues, iguales; si $m = n$, veremos en el párrafo siguiente que en general $AB \neq BA$ (es, por otra parte, evidente considerando

A y B, matrices cuadradas del mismo orden, como matrices de endomorfismos de un espacio vectorial E de dimensión finita).

4. Las propiedades de la transposición de una aplicación lineal y el teorema (§ 155, *b*)

$$M({}^tf) = {}^t[M(f)]$$

nos dan

$$^t(AB) = {}^tB\,{}^tA.$$

EJEMPLOS Y EJERCICIOS

1. Tomemos las notaciones del § 154 (ej. 1) y pongamos

$$X = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^m \end{pmatrix} = M(x, (a_i)), \qquad Y = \begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^n \end{pmatrix} = M(y, (b_j))$$

X es la matriz (unicolumna) asociada al vector x (en la base (a_i) de E) las fórmulas (1′) o (1″) del § 155, *a*, muestran que

$$M(f(x), (b_j)) = M(f, (a_i), (b_j))\, M(x, (a_i))$$

es decir,

$$Y = AX.$$

OBSERVACION

En ciertas obras se escribe una fórmula arriesgada $y = Ax$ que reemplaza a la vez la fórmula intrínseca, $y = f(x)$ y la fórmula entre matrices $Y = AX$; esto no tiene importancia si se usan siempre las mismas bases, pero resulta desastroso cuando se cambian las bases (ver § 161).

2. Escribir análogamente la fórmula $x^* = {}^tf(y^*)$ con la ayuda de matrices.

3. Si $A = (\alpha_i^j)$ es una matriz de tipo (m, n), m y n distintos de 1, $L = (\lambda_i)$ una matriz de una fila y $C = (\mu^j)$ una matriz de una columna, ¿en qué casos las matrices LA, AC están definidas? Calcular entonces estas matrices. Calcular LC cuando esté definido este producto.

4. Sean $A = (\alpha_i^j)$ una matriz de tipo (m, n) y D_1, D_2 dos matrices diagonales, ¿en qué casos los productos D_1A y AD_2 están definidos? Explicar los resultados obtenidos. Calcular D_1D_2 cuando este producto esté definido.

5. Si D y D′ son dos rectas del plano definidas por $(Ox, D) \equiv \alpha(\pi)$, $(Ox, D') \equiv \alpha'(\pi)$, S y S′ las matrices asociadas a las simetrías planas respecto a D y D′ (ver § 155, ej. 4), calcular S′S y SS′; ¿cuáles son las aplicaciones lineales asociadas a estas matrices? ¿En qué caso se tiene $S'S = SS'$?

6. Sea A, B, C tres matrices. Determinar sus tipos para que AB y (AB)C estén definidas. Verificar que en este caso BC y A(BC) están definidas.

7. Demostrar $^t(AB) = {}^tB\,{}^tA$, utilizando la regla de multiplicación de matrices.

8. Siendo A una matriz de tipo (m, n), demostrar que $A\,{}^tA$ y ${}^tA\,A$ son matrices cuadradas simétricas.

158. Anillo y álgebra $M_n(K)$. Grupo de las matrices cuadradas regulares de orden n

a) Anillo y álgebra $M_n(K)$

El conjunto de las matrices sobre K, cuadradas de orden n, representado $M_n(K)$, puede ser considerado como el conjunto de las matrices asociadas a los endomorfismos de un espacio vectorial E de dimensión n sobre K; la biyección $\mathcal{L}_K(E)$ sobre $M_n(K)$ definida por

$$f \to A = M(f, (a_i))$$

nos permite definir sobre $M_n(K)$ dos operaciones internas $A + B$, AB y una operación externa λA. La biyección recordada anteriormente y la estructura de anillo de $\mathcal{L}_K(E)$ nos muestran que $M_n(K)$, provisto de las operaciones $A + B$ y AB, tiene una estructura de *anillo* (observación 1 sobre el teorema 2 del § 96) y que estos dos anillos son isomorfos. Puesto que $\mathcal{L}_K(E)$ admite id_E por elemento unidad, $M_n(K)$ admite un elemento unidad

$$M(id_E, (a_i)) = I_n$$

también se justifica la definición dada en el § 154, *c*).

Este anillo no es conmutativo, como $\mathcal{L}_K(E)$, y como se puede ver en los ejemplos

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

(ver igualmente § 157, ej. 5).

Siempre igual que $\mathcal{L}_K(E)$, el anillo $M_n(K)$ está provisto de divisores de cero. Si no se quiere volver a los ejemplos dados en los ejercicios 2 y 3 del § 146, se puede simplemente constatar

$$\begin{pmatrix} \alpha & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & \beta \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{con} \quad \alpha \neq 0 \quad \beta \neq 0.$$

Igualmente $M_n(K)$ provisto de las tres operaciones $A + B$, AB, λA es una álgebra sobre K isomorfo a la álgebra $\mathcal{L}_K(E)$, de donde:

TEOREMA. — *El conjunto* $M_n(K)$ *de las matrices sobre* K, *cuadradas de orden* n, *provisto de las operaciones* $A + B$, AB *es un anillo, unitario, no conmutativo, que posee divisores de cero, isomorfo al anillo* $\mathcal{L}_K(E)$ $(\dim_K E = n)$; *provisto de las operaciones* $A + B$, AB, λA, $M_n(K)$ *es una álgebra sobre* K *isomorfa a la álgebra* $\mathcal{L}_K(E)$, $(\dim_K E = n)$.

b) Matrices escalares

En el anillo $M_n(K)$ se puede buscar las matrices $A = (\alpha_i^j)$ que permutan con toda matriz $M = (\mu_i^j)$, cuadrada de orden n, es decir, tales que $AM = MA$. Se tendrá para todo par (i, j)

$$\sum_{k=1}^{n} \alpha_k^j \mu_i^k = \sum_{l=1}^{n} \mu_l^j \alpha_i^l$$

igualdad válida para *toda* familia doble (μ^i_m) de elementos de K. Supongamos $i \neq j$ y consideremos la familia en que $\mu^j_i = 1$ con los restantes μ *todos* nulos, obtenemos $\alpha^j_i = 0$.

Suponiendo siempre $i \neq j$, consideremos la familia en que $\mu^j_i = 1$, y *todos* demás μ nulos, encontramos $\alpha^j_j = \alpha^i_i$.

Encontramos, pues, la condición necesaria (δ^j_i, índice de Kronecker, $\alpha \in$ K)

$$\alpha^j_i = \delta^j_i \alpha \Rightarrow A = \alpha I_n;$$

la condición es evidentemente suficiente.

Consideremos la aplicación de K en el conjunto K′ de todas las matrices de la forma αI_n ($\alpha \in$ K) definida por

$$\alpha \to \alpha I_n,$$

se tiene, cualesquiera que sean α, β, λ de K,

$$\begin{aligned} \alpha + \beta &\to (\alpha + \beta) I_n = \alpha I_n + \beta I_n \\ \alpha \beta &\to (\alpha \beta) I_n = (\alpha I_n)(\beta I_n) \\ \lambda \alpha &\to (\lambda \alpha) I_n = \lambda(\alpha I_n) \end{aligned}$$

esta aplicación biyectiva es, pues, un isomorfismo de K, considerado como álgebra sobre K, sobre K′ que tiene una estructura de álgebra sobre K: es una subálgebra de la álgebra $M_n(K)$, de donde:

Teorema y definición. — *En* $M_n(K)$ *las únicas matrices que permutan con toda matriz de* $M_n(K)$ *son las de la forma* αI_n ($\alpha \in$ K). *Estas matrices describen una subálgebra de* $M_n(K)$ *isomorfa a* K; *se llaman* matrices escalares.

Este isomorfismo justifica la denominación de matrices escalares que habíamos ya dado en el § 154, *c*). Por otra parte, K′ es isomorfo a $\mathcal{H}(E)$ (ver § 146, ej. 5, y § 147, ej. 8).

OBSERVACION

Si se considera únicamente las matrices cuadradas de orden n, no hay ningún inconveniente en representar α y αI_n con el mismo símbolo, lo que equivale a representar por 1 a la matriz I_n. Por el contrario, si se consideran matrices de orden cualquiera esta identificación puede conducirnos a errores, ya que naturalmente $m \neq n$ implica $I_m \neq I_n$. Algunos autores emplean la notación 1_n por I_n.

c) Matrices cuadradas inversibles (o regulares)

El isomorfismo $f \to M(f, (a_i))$ de $\mathcal{L}_K(E)$ sobre $M_n(K)$ ($\dim_K E = n$) demuestra que:

Teorema. — *Si* f *es un endomorfismo de un espacio vectorial* E *de dimensión* n *sobre* K, $M(f, (a_i))$ *es inversible si y sólo si* f *es un automorfismo.*

En este caso

$$M(f^{-1}) = [M(f)]^{-1}.$$

En efecto, la relación $f \circ f^{-1} = f^{-1} \circ f = \mathrm{id}_E$ nos da poniendo $A = M(f)$

$$AA^{-1} = A^{-1}A = I_n.$$

Por otra parte, el teorema del § 146, *b*) y la biyección $f \to M(f, (a_i))$ demuestran que:

TEOREMA. — *El conjunto de las matrices inversibles de* $M_n(K)$ *es un grupo (no conmutativo) isomorfo a* $\mathbf{GL}_n(K)$.

Naturalmente si dos matrices cuadradas A y B de orden n son inversibles, AB es inversible y

$$(AB)^{-1} = B^{1-}A^{-1}.$$

EJERCICIOS

1. Demostrar que el conjunto $\mathfrak{D}$ de las matrices K de orden n y diagonales es un subanillo (resp. una subálgebra) del anillo $M_n(K)$ (resp. de la álgebra $M_n(K)$), isomorfo al anillo (resp. a la álgebra) K^n (ver § 147).
2. Demostrar que el conjunto $\mathcal{C}_s$ de las matrices sobre K, triangulares superiores de orden n, es un subanillo (y una subálgebra) del anillo (o de la álgebra) $M_n(K)$. Igualmente $\mathcal{C}_i$ (ver § 156, ej. 2).
3. Demostrar que si A es una matriz inversible de orden n, A^p, cuando p describe $\mathbf{Z}$ (con $A^0 = I_n$) describe un subgrupo del grupo de las matrices inversibles de orden n, ¿es isomorfo al grupo aditivo $\mathbf{Z}$?
4. Si a y b son dos números reales cualesquiera, demostrar que el conjunto de las matrices reales cuadradas de orden 2 de la forma

$$M(a, b) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

provisto de la adición y de la multiplicación de las matrices es un cuerpo isomorfo a $\mathbf{C}$ (se tiene aquí un ejemplo de subanillo del anillo $M_2(\mathbf{R})$) teniendo una estructura de cuerpo conmutativo.
5. Si A es una matriz cuadrada de orden n inversible, demostrar que ${}^tA$ es inversible y que $({}^tA)^{-1} = {}^t(A^{-1})$.
6. Calcular A^n $(n \in \mathbf{N})$ para las matrices siguientes

$$\begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \qquad \begin{pmatrix} \operatorname{ch}\varphi & \operatorname{sh}\varphi \\ \operatorname{sh}\varphi & \operatorname{ch}\varphi \end{pmatrix} \qquad \begin{pmatrix} \operatorname{sh}\varphi & \operatorname{ch}\varphi \\ \operatorname{ch}\varphi & \operatorname{sh}\varphi \end{pmatrix}$$

Demostrar que estas matrices son inversibles, calcular A^{-1} y A^n $(n \in \mathbf{Z})$.
7. Si D es una matriz diagonal de orden n tal que para todo i de $[1, n]$, $\alpha_i^i = \lambda_i \neq 0$, demostrar que D es inversible, calcular D^n para todo entero racional n.
8. Demostrar que una matriz triangular es inversible si y sólo si sus elementos diagonales son todos no nulos.
9. Toda matriz cuadrada de orden n que verifica la ecuación

$$a_0 I_n + a_1 A^1 + \dots + a_n A^n = 0$$

$(a_0, a_1, \dots, a_n$ elementos de K, $a_0 \neq 0)$ es inversible. ¿Cuál es su inversa?
10. Siendo A y B dos matrices cuadradas de orden n tales que $AB = I_n$, demostrar que son inversibles y que $B = A^{-1}$.

159. Cálculos con matrices cuyos elementos pertenecen a un anillo

Sea E un conjunto provisto de una adición y dos matrices de tipo (m, n) de elementos en E, A y B, pondremos por definición

$$(1) \qquad A + B = (\alpha_i^j) + (\beta_i^j) = (\alpha_i^j + \beta_i^j) \qquad {\scriptstyle (1 \leq i \leq m) \atop (1 \leq j \leq n)}$$

Si E está provisto de una multiplicación pondremos $(\lambda \in E)$

$$(2) \qquad \lambda A = \lambda(\alpha_i^j) = (\lambda\alpha_i^j).$$

Finalmente si E posee una adición y una multiplicación, dadas las dos matrices

$$A = (\alpha_i^j) \in M_E(n, m), \qquad B = (\beta_i^j) \in M_E(p, n)$$

pondremos por definición

$$(3) \qquad AB = (\gamma_i^j) \qquad \gamma_i^j = \sum_{k=1}^{n} \alpha_k^j \beta_i^k.$$

Naturalmente las propiedades de las operaciones definidas sobre las matrices de elementos en un cuerpo conmutativo no se extienden, en general, a las operaciones (1), (2) y (3) definidas más arriba cuando no se sabe nada de las propiedades de la adición y de la multiplicación en E: ello nos dice el poco interés de estas definiciones (1), (2) y (3) en el caso general.

Pero supongamos que E sea un *anillo conmutativo unitario* A.

Desde luego *si* A *es un anillo íntegro, se le puede considerar como un* subanillo de su cuerpo de las fracciones K; luego las operaciones (1), (2) y (3) poseen todas las propiedades estudiadas en los párrafos precedentes. Notemos, sin embargo, que una matriz cuadrada de $M_n(A)$ *puede ser inversible en* $M_n(K)$ y no serlo en $M_n(A)$ (ver ej. 215, fin de este capítulo, y ej. 273, del capítulo 9).

Si A es un anillo conmutativo unitario se demostrará, a título de ejercicio, que las propiedades demostradas para las matrices de elementos en un cuerpo conmutativo, se aplican a las matrices de elementos en A, reemplazando las nociones de espacio vectorial de dimensión finita sobre K por la de A-módulo libre de tipo finito (ver capítulo 7, ej. 179).

Así, a todo homomorfismo f de M en N, donde M y N son A-módulos expresados en bases (a_i) $(1 \leq i \leq m)$ y (b_j) $(1 \leq j \leq n)$ viene asociada una matriz de $M_A(m, n)$ y la aplicación definida por

$$f \to M(f, (a_i), (b_j))$$

es una biyección de $\mathfrak{L}_A(M, N)$ sobre $M_A(m, n)$; se deduce que estos dos A-módulos son isomorfos y asimismo los anillos $\mathfrak{L}_A(M)$, si M es un A-módulo que tiene una base de n elementos, y $M_n(A)$.

III. Cambio de base

160. Acción de un cambio de bases sobre las coordenadas de un vector de un espacio vectorial E

Sea E un espacio vectorial de dimensión n sobre K, $B = (a_i)$ una base de E y $B' = (a'_i)$ una segunda base de E; consideremos la matriz $P = (p^j_i)$ que tiene por columna i las coordenadas de a'_i referidas a B tenemos

$$a'_i = \sum_{j=1}^{n} p^j_i a_j$$

Ahora bien, en el § 155, hemos visto que

$$f(a_i) = \sum_{j=1}^{r} \alpha^j_i b_j \Leftrightarrow A = \alpha^j_i = M(f,\ (a_i),\ (b_j))$$

Deducimos de ello que si id_E representa la aplicación idéntica de E, las relaciones

$$id_E(a'_i) = a'_i = \sum_{j=1}^{n} p^j_i a_j \qquad (i = 1,\ 2,\ \ldots,\ n)$$

son equivalentes a

(1) $$P = (p^j_i) = M(id_E.\ (a'_i),\ (a_i))$$

(ATENCIÓN: E *espacio de salida* está referido a (a'_i) y E *espacio de llegada* está referido a (a_i).)

Luego P matriz asociada a id_E es inversible, la fórmula (ver § 157, *a*)

$$M(id_E,\ (a'_i),\ (a_i))M(id_E,\ (a_i),\ (a'_i)) = M(id_E,\ (a_i),\ (a_i)) = I_n$$

demuestra que

(1') $$M(id_E,\ (a_i)\ \ (a'_i)) = P^{-1}.$$

Luego si se pone

(2) $$x = \sum_{i=1}^{n} x^i a_i = \sum_{i=1}^{n} x'^i a'_i$$

tendremos según (1) (ver § 155)

(3) $$x^j = \sum_{i=1}^{n} p^j_i x'^i \qquad (1 \leq j \leq n)$$

y según (2) poniendo $P^{-1} = (q^j_i)$

$$(4) \qquad x'^j = \sum_{i=1}^{n} q^j_i x^i \qquad (1 \leq j \leq n)$$

que se puede escribir en forma matricial poniendo

$$X = M(x, (a_i)), \qquad X' = M(x, (a'_i)) \qquad (\text{ver § 154, ej. 1})$$
$$X = PX' \qquad X' = P^{-1}X$$

TEOREMA Y DEFINICIÓN. — *Si* (a_i) y (a'_i) $(1 \leq i \leq n)$ *son dos bases de un espacio vectorial* E *de dimensión* n *sobre* K, *la matriz* P *que tiene por columna* i*-ésima las coordenadas de* a'_i *respecto a la base* (a_i) *es inversible; se le llama la* matriz de cambio *de la base* (a_i) *a la base* (a'_i)*; además.*

$$P = M(\mathrm{id}_E, (a'_i), (a_i)), \qquad P^{-1} = M(\mathrm{id}_E, (a_i), (a'_i))$$
$$X = PX', \qquad X' = P^{-1}X$$

donde X *y* X′ *son las matrices unicolumnas asociadas a un vector* x *de* E, X *en la base* (a_i), X′ *en la base* (a'_i).

OBSERVACIONES

Cuando se pasa de una base (a_i) a otra base (a'_i) se dice algunas veces que (a_i) es la *base «antigua»*, y (a'_i) la *«nueva»*, $(x^1, x^2, \ldots, x^n)$ las coordenadas *«antiguas»* de x, $(x'^1, x'^2, \ldots, x'^n)$ las *«nuevas»*. Se ve que la matriz de paso P da las *«coordenadas antiguas»* en función de las *«nuevas»*: lo que en general es ventajoso, pues si se cambia de coordenadas en un problema se tiene las relaciones entre las coordenadas antiguas, por ejemplo, $F(x^1, x^2, \ldots, x^n) = 0$, para tener las relaciones correspondientes entre las nuevas coordenadas, es cómodo tener las fórmulas dando las x^j en función de las x'^i (y no al contrario).

2. Para estar seguro de no equivocarse, se puede hacer el cálculo siguiente que es fácil, pero que tiene el inconveniente de que no nos da la verdadera naturaleza de la matriz de paso P. Las fórmulas

$$a'_i = \sum_{j=1}^{n} p^j_i a_j \qquad x = \sum_{i=1}^{n} x^i a_i = \sum_{i=1}^{n} x'^i a'_i$$

dan

$$x = \sum_i x'^i \sum_j p^j_i a_j = \sum_i \sum_j p^j_i x'^i a_j = \sum_j a_j \sum_i p^j_i x'^i = \sum_j a_j x^j$$

de donde, siendo única la descomposición de x sobre la base (a_i),

$$x^j = \sum_i p^j_i x'^i.$$

EJERCICIO

Con la misma notación que antes, se designa por (a^{*i}) la base dual de (a_i) y (a'^{*i}) la base dual de (a'_i), demostrar que la matriz de paso de (a^{*i}) a (a'^{*i}) es $({}^tP)^{-1}$ y que $Y'^* = Y^*P$, Y^* es la matriz de una fila que da las coordenadas de una forma y^* de E^* respecto a la base (a^{*i}) e (Y'^*) la matriz de una fila dando las coordenadas de la misma forma en la base (a'^{*i}).

161. Acción del cambio de bases en E y en F en una matriz de una aplicación lineal de E en F

a) Sea f una aplicación lineal de E en F con $\dim_K E = m$, $\dim_K F = n$. Sean (a_i) y (a'_i) dos bases de E y (b_j) y (b'_j) dos bases de F. Pongamos

$$A = M(f, (a_i), (b_j)) \qquad A' = M(f, (a'_i), (b'_j))$$
$$P = M(\mathrm{id}_E, (a'_i), (a_i)) \qquad Q = M(\mathrm{id}_F, (b'_j), (b_j)).$$

Consideremos el diagrama

$$\underset{(a'_i)}{E} \xrightarrow{\mathrm{id}_E} \underset{(a_i)}{E} \xrightarrow{f} \underset{(b_j)}{F} \xrightarrow{\mathrm{id}_F} \underset{(b'_j)}{F}$$

tenemos

$$f = \mathrm{id}_F \circ f \circ \mathrm{id}_E$$

de donde aplicando la fórmula del § 157, *a*)

$$M(f, (a'_i), (b'_j)) = M(f, (a_i), (b_j))\ M(\mathrm{id}_F, (b_j), (b'_j))\ M(f, \mathrm{id}_E, (a'_i), (a_i))$$
$$A' = Q^{-1}AP.$$

OBSERVACION

Se puede también operar únicamente sobre las matrices. En efecto, con la notación del párrafo precedente

$$Y = AX, \qquad X = PX', \qquad Y = QY'$$

dan

$$QY' = APX'$$

de donde multiplicando por la izquierda por Q^{-1}

$$Y' = Q^{-1}APX' = A'X'$$

es decir, $(Q^{-1}AP - A')X' = 0$ *cualquiera que sea* X', aplicando esta fórmula a las matrices asociadas a los m vectores a'_i en la base (a'_i) se encuentra fácilmente

$$A' = Q^{-1}AP.$$

Teorema. — *Dados dos espacios vectoriales* E, F *de dimensiones finitas sobre* K, (a_i) *y* (a'_i) *dos bases de* E, (b_j) *y* (b'_j) *dos bases de* F, P *la matriz de paso de* (a_i) *a* (a'_i) *y* Q *la matriz de paso de* (b_j) *a* (b'_j), *para toda aplicación lineal* f *de* E *en* F *es*

$$A' = Q^{-1}AP$$

donde A *es la matriz asociada a f respecto a las bases* (a_i) *y* (b_j) *y* A′ *respecto a las bases* (a'_i), (b'_j).

Corolario. — *Para todo endomorfismo f de* E

$$A' = P^{-1}AP$$

donde A *y* A′ *son las matrices asociadas a f, respectivamente, respecto a la base* (a_i) *y a la base* (a'_i) *y* P *la matriz de paso de* (a_i) *a* (a'_i).

b) Matrices equivalentes

Sea A y B dos matrices de tipo (m, n) tales que existen dos matrices cuadradas inversibles R y S que verifican la relación

$$B = R\,A\,S$$

se observa que esta relación binaria definida sobre $M_K(m, n)$ es una relación de equivalencia; en efecto, $R \in GL_n(K)$ y $S \in GL_m(K)$. Por tanto:

1. Para toda matriz A de $M_K(m, n)$,

$$A = I_n\,A\,I_m.$$

2. $B = R\,A\,S \Rightarrow A = R^{-1}\,B\,S^{-1}$.
3. $B = R\,A\,S$ y $C = T\,B\,U$ implica que

$$C = T(R\,A\,S)\,U = (T\,R)\,A\,(S\,U)$$

con T R y S U matrices inversibles de órdenes respectivas n y m; de donde:

Teorema y definición. — *La relación binaria entre matrices de* $M_K(m, n)$: *«existe* R *inversible de* $M_n(K)$ *y* S *inversible de* $M_m(K)$ *tales que* B = RAS*» es una relación de equivalencia. Las matrices* A *y* B *se llaman matrices* equivalentes.

Se ve inmediatamente (ver *a*), anterior) que dos matrices de tipo (m, n) son equivalentes si y sólo si son asociadas a la misma aplicación lineal de E, de dimensión m, en F de dimensión n: es suficiente poner $P = S$ y $Q = R^{-1}$.

c) Matrices semejantes

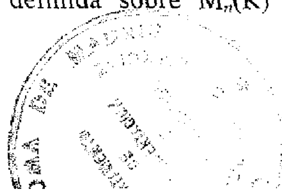
Definición. — *Dos matrices cuadradas* A *y* B *de orden n son semejantes si existe una matriz cuadrada, de orden n, inversible,* P *tal que*

$$B = P^{-1}AP.$$

Se ve inmediatamente que esta relación binaria definida sobre $M_n(K)$ es una relación de equivalencia; en efecto:

1. Para toda matriz A de $M_n(K)$,

$$A = (I_n)^{-1}\,A\,I_n.$$



2. $B = P^{-1}AP \Rightarrow A = (P^{-1})^{-1}BP^{-1}$.
3. $B = P^{-1}AP$ y $C = Q^{-1}AQ$ implica

$$C = P^{-1}Q^{-1}AQP = (QP)^{-1}A(QP)$$

pues al ser P y Q inversibles, QP lo es también.

La definición anterior y los corolarios del subpárrafo *a*), demuestran que dos matrices A y B cuadradas de orden n, son semejantes si y sólo si están asociadas a un mismo endomorfismo de E, de dimensión n.

Uno de los fines de alguno de los próximos capítulos, el capítulo 13, será el de, dado un endomorfismo f de E, encontrar una base de E de manera que la matriz asociada a f en esta base sea la más "simple" posible, es decir, dada una matriz A de $M_n(K)$, encontrar una matriz semejante a A lo más "simple" posible: ya veremos en este capítulo cómo hay que entender esta noción de simplicidad.

162. Rango de una matriz

a) DEFINICIÓN. — *Dada una matriz* A *sobre* K *de tipo* (m, n) *se llama* rango *de la matriz* A *y se designa* rg(A), *el rango del sistema de sus vectores columnas (en* K^n*).*

Según las definiciones del § 138, rg (A) es, por tanto, el número máximo de columnas de A linealmente independientes. Si se considera A como matriz de una aplicación lineal f de E en F (dim E $= m$, dim F $= n$).

$$A = M(f \quad (a_i), \ (b_j))$$

las columnas de A representan los m vectores $f(a_i)$ de F, su rango es, en consecuencia, el rango de f (§ 143, teorema 7), luego rg (A) $\leq$ inf (m, n).

Por otra parte, ${}^tA = M({}^tf, \ (b^{*j}), \ (a^{*i}))$, luego (§ 155, *b*)

$$\text{rg}\ ({}^tA) = \text{rg}\ ({}^tf) = \text{rg}\ (f) = \text{rg}\ (A)$$

de donde:

TEOREMA. — *Si* A *es la matriz de tipo* (m, n) *asociada a una aplicación lineal* f *de* E *en* F, *los cuatro números siguientes son iguales:*

1. *Rango de los vectores columnas (en* K^n*).*
2. *Rango de los vectores filas (en* K^m*).*
3. *Rango de* f.
4. *Rango de* ${}^tf$.

b) Caracterización de las matrices de tipo (m, n) de rango r

Sea A una matriz de tipo (m, n) asociada a una aplicación lineal f de E de dimensión m en F de dimensión n. Escojamos en E y F las bases (a_i), (b_j), como se ha dicho en el § 143 (corolario 3) y § 155 (corolario 1), tenemos

$$M(f, (a_i), (b_j)) = \left(\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right) \begin{array}{l} \updownarrow r \\ \updownarrow n-r \end{array}$$
$$\qquad\qquad\qquad\qquad \overleftrightarrow{\ r\ } \quad \overleftrightarrow{\ m-r\ }$$

TEOREMA. — *Todas las matrices de tipo* (m, n) *de rango* f *son equivalentes a la matriz de tipo* (m, n)

$$\left(\begin{array}{c|c} I_r & O \\ \hline O & O \end{array}\right)$$

llamada matriz canónica de rango r *de tipo* (m, n).

COROLARIO 1. — *Para que dos matrices de tipo* (m, n) *sean equivalentes, es necesario y suficiente que sean del mismo rango.*

Si son equivalentes, las dos matrices representan respecto a las bases diferentes la misma aplicación lineal f: son, pues, del mismo rango, el de f. Si tienen el mismo rango, sea r, son equivalentes a la matriz canónica de tipo (m, n) de rango r y, en consecuencia, son equivalentes entre sí.

COROLARIO 2. — *Para que una matriz cuadrada de orden* n *sobre el cuerpo conmutativo* K *sea inversible, es necesario y suficiente que sea de rango* n.

Esta matriz representa, en efecto, un endomorfismo f de E, de dimensión n sobre K, y $\text{rg}\,(f) = n = \dim \text{E}$ implica que f es un automorfismo (ver § 143, *b*).

c) Determinación del rango de una matriz de tipo (m, n)

Es suficiente operar sobre los vectores columnas de A como lo hemos hecho para un sistema de vectores en el § 138. Esta "manipulación" sobre las columnas puede interpretarse como una sucesión de cambios de bases en E, estando A asociada a una aplicación lineal de E en F, después de un número finito de cambios de base en E se encuentra una matriz A′ equivalente a A de la forma A′ (ver ej. 230 fin del capítulo)

$$A' = \left(\begin{array}{ccccc|c} \beta_1^1 & 0 & \cdots & \cdots & 0 & \\ \cdot & \beta_2^2 & & & 0 & \\ \cdot & \cdot & \cdot & & \cdot & \\ \cdot & \cdot & & \cdot & \cdot & O \\ \cdot & \cdot & & \cdot & 0 & \\ \cdot & \cdot & & \cdot & \beta_r^r & \\ \hline \cdot & \cdot & & & \cdot & \\ \cdot & \cdot & & & \cdot & O \\ \beta_1^n & \beta_2^n & \cdots & \cdots & \cdot & \end{array}\right) \begin{array}{l} \updownarrow r \\ \\ \updownarrow n-r \end{array}$$

con $\beta_i^i \neq 0$ $(1 \leq i \leq r)$, o los dos casos particulares siguientes

$$\left(\begin{array}{cccc|c} \beta_1^1 & 0 & \cdots & 0 & \\ \vdots & \beta_2^2 & \ddots & \vdots & \\ \vdots & \vdots & \ddots & 0 & O \\ \vdots & \vdots & & \vdots & \\ \beta_1^n & \beta_2^n & \cdots & \beta_n^n & \end{array}\right) \qquad \begin{pmatrix} \beta_1^1 & 0 & \cdots & 0 \\ \vdots & \beta_2^2 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \vdots & \vdots & & \beta_m^m \\ \vdots & \vdots & & \vdots \\ \beta_1^n & \beta_2^n & \cdots & \beta_m^n \end{pmatrix}$$

$$n = r \qquad\qquad\qquad\qquad m = r$$

el rango de estas tres matrices es evidentemente r ($r = n$ en el segundo caso, f es suprayectiva, $r = m$ en el tercer caso, f es inyectiva).

Se puede también razonar sobre los vectores filas de A, esta "manipulación sobre las filas" puede interpretarse como una sucesión de cambios de bases en F: se encuentra matrices equivalentes a A análogas a las transpuestas de las tres matrices encontradas más arriba (ver ej. 231, fin del capítulo).

Ejercicios

NOTA: Salvo mención contraria los elementos de las matrices se tomarán en un cuerpo conmutativo K. En los ejercicios de inversión de matrices se procurará no utilizar los determinantes.

181. En el conjunto $\mathbf{M}_n(\mathrm{K})$ se considera el subconjunto E descrito por las matrices $\mathrm{M}(a, b)$ en que todos los elementos de la diagonal principal son iguales a a, y todos los demás iguales a b; demostrar que cuando a y b describen **K**, E tiene una estructura de espacio vectorial sobre **K**. Indicar una base. ¿Cuál es la dimensión de E?

182. Resolver el ejercicio precedente para el conjunto F de las matrices $\mathrm{M}(a_1, \ldots, a_n)$ cuya i-ésima línea es

$$a_1a_2 \ldots a_{i-1}a_ia_i \ldots a_i$$

cuando $(a_1, \ldots, a_n)$ describe $\mathbf{K}^n$.

183. Se dice que una matriz M de $\mathbf{M}_3(\mathbf{R})$ es mágica si las ocho sumas

$$(j = 1, 2, 3) \sum_i a_{ij}, \qquad (i = 1, 2, 3) \sum_i a_{ij}, \qquad \sum_j a_{ij}, \qquad a_{13} + a_{22} + a_{31}$$

son iguales; se designa por s el valor común de esas ocho sumas y por $\mathrm{M}(s)$ una de las matrices correspondientes.

a) Demostrar que el conjunto de las matrices mágicas de $\mathbf{M}_3(\mathbf{R})$ es un subespacio vectorial del espacio vectorial $\mathbf{M}_3(\mathbf{R})$.

b) ¿Cuál es el valor de s si $\mathrm{M}(s)$ es antisimétrica? Construir todas las matrices mágicas antisimétricas.

c) Construir todas las matrices mágicas simétricas (se construirá primero las correspondientes a $s = 0$).

d) ¿Cuál es la dimensión del subespacio vectorial de $\mathbf{M}_3(\mathbf{R})$ descrito por las matrices mágicas?

M.G.P. (extracto) y *Concours de Mines* (extracto).

184. Se considerarán las matrices

$$\mathrm{M}(s) = \begin{pmatrix} s & 0 \\ 0 & 1/s \end{pmatrix}, \qquad \mathrm{N}(t) = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}, \qquad \mathrm{P}(u) = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}$$

donde s, t, u son los elementos de un cuerpo K dado: se supone $s \neq 0$. Sea

$$\mathrm{X} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

una matriz de coeficientes en K. Demostrar que, para que pueda ponerse X bajo la forma

(1) $$\mathrm{X} = \mathrm{M}(s)\mathrm{N}(t)\mathrm{P}(u)$$

es necesario y suficiente que se tenga $a \neq 0$, $ad - bc = 1$; demostrar que la descomposición (1) es entonces única.

Sea

$$\mathrm{W} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

demostrar que si $a = 0$, $ad - bc = 1$, se puede poner X en la forma

(2) $$X = M(s)N(t)W.$$ (M.G.P.)

185. En $M_2(\mathbf{C})$ se considera las matrices

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

calcular las matrices

$$YZ - ZY, \quad ZX - XZ, \quad XY - YX, \quad X^2 + Y^2 + Z^2.$$

186. En $M_2(\mathbf{K})$ encontrar todas las matrices que permutan con $A = \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix}$; demostrar que estas matrices describen una subálgebra de $M_2(K)$.

187. Siendo A una matriz de $M_2(\mathbf{K})$, demostrar que existe α y β de $\mathbf{K}$ tales que: $A^2 - \alpha A + \beta I_2 = 0$. ¿Cuál es la inversa de A si A es inversible?

188. Siendo A una matriz *no escalar* de $M_2(\mathbf{K})$, encontrar todas las matrices de $M_2(\mathbf{K})$ permutando con A; demostrar que son de la forma $\lambda I_2 + \mu A$ (utilizar el ejercicio 187).

189. Sea A una matriz triangular de $M_n(\mathbf{K})$ en la que todos los elementos diagonales son nulos. Demostrar que $A^n = 0$ (si A es triangular superior y si $A = M(f, (a_i))$ se observará que $f(a_1) = 0$ y que para $k > 1$, $f(a_k)$ pertenece al subespacio engendrado por $a_1, a_2, \ldots, a_{k-1}$).

190. Hallar todas las matrices A de $M_2(\mathbf{K})$ tales que $A \neq 0$ y $A^2 = 0$.

191. Sea f un endomorfismo de E de dimensión n sobre K tal que $f \neq 0$ y $f^2 = 0$. Se pone $r = \mathbf{rg}(f)$.

a) Demostrar que $\operatorname{Im} f \subset \operatorname{Ker} f$; deducir de ello que $2r \leq n$; sea F un suplementario de Ker f, F es de dimensión r; siendo (a_i) $(1 \leq i \leq r)$ una base de F, demostrar que $b_i = f(a_i)$ son independientes y que existe $n - 2r$ vectores de Ker f, $c_1, \ldots, c_{n-2r}$, tales que

$$\{a_1, \ldots, a_r, b_1, \ldots, b_r, c_1, \ldots, c_{n-2r}\}$$

sea una base de E.

b) ¿Cuál es la matriz de f en la base precedente?

c) ¿Cómo se simplifican los resultados si $n = 2r$? (V. cap. 7, ej. 162).

192. Se dice que un endomorfismo f de un espacio vectorial E de dimensión n sobre K es *nilpotente* de índice p, si existe $p > 1$ tal que $f^{p-1} \neq 0$, $f^p = 0$, se dirá igualmente que una matriz A de $M_n(\mathbf{K})$ es *nilpotente* de índice p si existe $p > 1$ tal que $A^{p-1} \neq 0$ y $A^p = 0$.

a) Demostrar que si f es nilpotente y si existe λ de $\mathbf{K}$ y $x \neq 0$ de E tal que $f(x) = \lambda x$ entonces $\lambda = 0$.

b) Siendo f nilpotente de índice p, demostrar que si x es tal que $f^{p-1}(x) \neq 0$ los vectores

$$x = f^0(x), \quad f^1(x), \ldots, f^{p-1}(x)$$

son linealmente independientes.

Deducir que f es nilpotente de índice n, si y solamente si existe una base (a_i) de E, tal que para la matriz

$$A = M(f, (a_i)) = (\alpha_i^j)$$

todos los α_i^j son nulos salvo $\alpha_i^{i+1} = 1$ $(1 \leq i \leq n-1)$.

193. Si A es una matriz nilpotente de índice p de $M_n(\mathbf{K})$ (V. ej. 192), demostrar que $I_n - A$ es inversible y tiene como inversa

$$I_n + A + A^2 + \ldots + A^{p-1}.$$

194. Tenemos las matrices

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \qquad T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Demostrar que $T^3 = 0$. Deducir A^n para n entero natural (se observará que $A = I_3 + T$ y que I_3 permuta con T).

195. Buscar todas las matrices A de $M_2(\mathbf{R})$ tales que $A^2 = I_2$.

196. Si **K** es un cuerpo conmutativo de característica distinta de 2, se considera los endomorfismos f de un espacio vectorial E de dimensiones n sobre **K** tales que $f^2 = \mathrm{id}_E$. Se pone $g = \mathrm{id}_E + f$, $h = \mathrm{id}_E - f$.

a) Demostrar que $g(E)$ y $h(E)$ son estables por f y son dos subespacios suplementarios de E. ¿Cuáles son las aplicaciones inducidas por f, respectivamente, en $g(E)$ y $h(E)$?

b) Deducir de la pregunta anterior que toda matriz M de $M_n(\mathbf{K})$ tal que $M^2 = I_n$ es semejante a una matriz de la forma

$$\begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix} \qquad \text{con} \qquad p \in \mathbf{N}, \qquad q \in \mathbf{N}, \qquad p + q = n.$$

c) ¿Cuál es la forma general de una matriz de $M_n(\mathbf{K})$ tal que $M = I_n$?

197. Encontrar todas las matrices A de $M_2(\mathbf{R})$ tales que $A^2 = -I_2$.

198*. Sea E un espacio vectorial de dimensión finita sobre **R**. Se designa con f un endomorfismo de E tal que $f^2 = -\mathrm{id}_E$ y se define una aplicación de $\mathbf{C} \times E$ en E mediante

$$(a + ib, x) \rightarrow ax - bf(x) \qquad (a, b \in \mathbf{R}, x \in E).$$

a) Demostrar que la aplicación precedente permite dar al conjunto E una estructura de espacio vectorial sobre **C**: se designará por E′ el conjunto E provisto de esta estructura.

b) Calcular dim E en función de dim E′. Deducir que si existe f endomorfismo de E tal que $f^2 = -\mathrm{id}_E$, E es de dimensión par $2n$.

c) Demostrar que existe entonces una base de E descrita por $a_1, \ldots, a_n, f(a_1), \ldots, f(a_n)$. Buscar la matriz de f respecto a esta base. Deducir que sobre todo espacio vectorial E de dimensión par sobre **R** existe al menos un endomorfismo tal que $f^2 = -\mathrm{id}_E$.

d) Sea g una aplicación de E sobre sí mismo, demostrar que las dos propiedades siguientes son equivalentes: α) g es un endomorfismo de E espacio vectorial sobre **R**, que permuta con f; β) g es un endomorfismo del espacio vectorial E′ sobre **C**.

e) Sea G una matriz de $M_n(\mathbf{C})$, se pone $G = A + iB$ donde A y B son elementos de $M_n(\mathbf{R})$ y

$$\varphi(G) = \begin{pmatrix} A & B \\ -B & A \end{pmatrix}$$

demostrar que φ es un homomorfismo del anillo $M_n(\mathbf{C})$ en el anillo $M_{2n}(\mathbf{R})$. Demostrar que $\mathrm{Im}(\varphi)$ está descrito por las matrices de $M_{2n}(\mathbf{R})$ que permutan con la matriz

$$\begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}.$$

199. Se tienen las matrices $M(a, b)$ definidas en el ejercicio 181.

a) Calcular la suma y el producto de dos matrices $M(a, b)$ y $M(a', b')$ (se introducirá I_n y $M(1, 1)$ y se utilizará el ejercicio 181).

b) Si (a, b) describe $\mathbf{K}^2$, demostrar que $M(a, b)$ describe un subanillo de $M_n(\mathbf{K})$. Demostrar que de los cálculos anteriores surge la definición de una estructura de anillo sobre el conjunto $\mathbf{K}^2$.

200. *Se da un número complejo no real* $a = r\ (\cos\theta + \mathrm{sen}\ \theta)$ que permanecerá fijo en todo el problema y se tienen las matrices

$$M(x, y) = \begin{pmatrix} x & y \\ -r^2 y & x + 2ry\cos\theta \end{pmatrix}.$$

Sea E el conjunto de estas matrices cuando (x, y) describe $\mathbf{R}^2$ y F este conjunto cuando (x, y) describe $\mathbf{C}^2$.

a) Demostrar que E es un subanillo de $M_2(\mathbf{R})$ teniendo una estructura de cuerpo conmutativo.

b) Demostrar que todo número complejo z puede escribirse de una manera única $z = x + ay$ $(x, y \in \mathbf{R})$. Se pone $M(z) = M(x, y)$ que se puede decir de la aplicación de **C** en E definida por

$$z \to M(z).$$

Calcular $[M(z)]^n$.

c) Demostrar que F es un subanillo de $M_2(\mathbf{C})$. ¿Es F íntegro? ¿Tiene una estructura de cuerpo?

(M.G.P.)

201. Si (x_n), (y_n) son dos sucesiones reales convergentes y de límites respectivos x e y cuando n tiende hacia $+\infty$ se dirá que la matriz

$$M(x_n, y_n) = \begin{pmatrix} x_n & y_n \\ -y_n & x_n \end{pmatrix}$$

tiene por límite $M(x, y) = \begin{pmatrix} x & y \\ -y & x \end{pmatrix}$ cuando n tiende hacia $+\infty$. Si α es un número real, encontrar el límite para n infinito de la matriz

$$\begin{pmatrix} 1 & \frac{\alpha}{n} \\ -\frac{\alpha}{n} & 1 \end{pmatrix}^n$$

(se pondrá $\frac{\alpha}{n} = \operatorname{tg}\ \varphi,\ -\pi/2 < \varphi < \pi/2$).

202. Sea E el subconjunto de $M_2(\mathbf{Q})$ descrito por las matrices

$$M(a,\ b) = \begin{pmatrix} a & b \\ 2b & a \end{pmatrix}$$

cuando a y b describen **Q**.

a) Demostrar que E es una subálgebra de $M_2(\mathbf{Q})$.

b) Demostrar que para la adición y la multiplicación E tiene una estructura de cuerpo isomorfo a $\mathbf{Q}[\sqrt{2}]$ (ver cap. 5, ej. 97).

c) Los resultados anteriores, ¿son los mismos si se reemplaza **Q** por **R**? (Se escribirá $M(a,\ b) = aI_2 + bJ$ y se calculará J^2.)

203. Sea E el subconjunto de $M_3(\mathbf{Q})$ descrito por las matrices

$$M(a,\ b,\ c) = \begin{pmatrix} a & b & c \\ 3c & a - 3c & b \\ 3b & -3b + 3c & a - 3c \end{pmatrix}$$

cuando a, b, c describen **Q**.

a) Demostrar que E es una subálgebra de $M_3(\mathbf{Q})$.

b) Demostrar que E tiene una estructura de cuerpo.

c) ¿Subsisten los resultados precedentes si se reemplaza **Q** por **R**?

(Se escribirá $M(a,\ b,\ c) = aI_3 + bJ + cK$, y se calculará, J^2, K^2, JK, KJ.)

204. Sea p una permutación de $[1,\ n]$, si E es un espacio vectorial sobre K referido a la base (e_i) $(1 \leq i \leq n)$ se considera el automorfismo f_p definido por

$$i = 1,\ 2,\ \ldots,\ n \qquad f_p(e_i) = e_{p(i)}.$$

Se dice que

$$A_p = M(f_p,\ (e_i)) = (a_i^j)$$

es una *matriz de permutación.*

Demostrar que

$$a_i^j = \delta_{p(i)}^j$$

(δ_k^j es el símbolo de KRONECKER). Deducir que

$$A_p A_q = A_{p \circ q}$$

y que el conjunto de las matrices A_p provisto de la multiplicación de las matrices es un grupo isomorfo al grupo simétrico $\mathfrak{S}_n$.

205. Se dice que una matriz de $M_n(\mathbf{Q})$ es *monomial* si en cada fila y en cada columna se encuentra un elemento y uno sólo no nulo. Demostrar que toda matriz monomial es el producto de una matriz de permutación (V. ej. 204) por una matriz diagonal; representar de modo explícito esta última, con a_i el único elemento no nulo de la matriz monomial situada en la columna i.

206. Se pone $a = \cos(2\pi/n) + i \operatorname{sen}(2\pi/n)$ y se designa por X e Y las matrices de $M_n(\mathbf{C})$ de elementos generales respectivos

$$x_{pq} = a^{(p-1)(q-1)}, \quad y_{pq} = a^{-(p-1)(q-1)}$$

(p índice de fila, q índice de columna). Calcular

$$X^2, \qquad Y^2, \qquad XY, \qquad YX.$$

¿Cuál es la inversa de X? M.G.P. (extracto).

207. Los elementos de las matrices cuadradas, reales, de orden 3, que consideramos son funciones reales de t real que tienen desarrollos limitados de orden 2 en una vecindad de cero; a cada matriz M se asocia la matriz M_1 obtenida reemplazando en M cada elemento por la parte regular de orden 2 de su desarrollo limitado. El espacio euclídeo $\mathbf{R}^3$ está referido a una referencia ortonormal, se consideran las matrices X e Y representando respectivamente una rotación de ángulo $\alpha(t)$ alrededor de Ox y una rotación de ángulo $\beta(t)$ alrededor de Oy. Se supone que $\alpha(t)$ y $\beta(t)$ son infinitésimos de orden 1 en la vecindad de $t = 0$.

a) Demostrar que

$$X = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\operatorname{sen}\alpha \\ 0 & \operatorname{sen}\alpha & \cos\alpha \end{pmatrix} \qquad X_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-\dfrac{\alpha^2}{2} & -\alpha \\ 0 & \alpha & 1-\dfrac{\alpha^2}{2} \end{pmatrix}.$$

b) Calcular Y e Y_1 (atención: una rotación de $\pi/2$ alrededor de Oy aplica Oz sobre Ox).

c) Se pone $A = YX$, $B = Y^{-1}X^{-1}$ (se observará que dos rotaciones de ángulo θ y $-\theta$ alrededor de un mismo eje son recíprocas una de la otra).
Calcular A_1 y B_1.

d) Calcular $Z_1 = (BA)_1$. Demostrar que Z_1 está asociada a una matriz Z que se interpretará mediante una rotación alrededor de Oz.

208. Se considera el subconjunto T de $M_2(\mathbf{R})$ descrito por

$$A = \begin{pmatrix} a & c \\ 0 & b \end{pmatrix}$$

se designa por T_+ el subconjunto de T descrito por las matrices A tales que $a > 0$, $b > 0$.

a) Calcular $A^2, A^3, \ldots, A^p$ y la matriz

$$B_n = \sum_{p=0}^{n} \frac{1}{p!} A^p = \begin{pmatrix} \alpha_n & \gamma_n \\ 0 & \beta_n \end{pmatrix}$$

demostrar que α_n, β_n, γ_n tienen límites α, β, γ que se calcularán, cuando n tiende hacia el infinito. Se escribirá

$$B = f(A) = \begin{pmatrix} \alpha & \gamma \\ 0 & \beta \end{pmatrix}$$

(se distinguirá los casos $a \neq b$ y $a = b$).

b) La aplicación f de T en T así definida, ¿es lineal? ¿es inyectiva? ¿es suprayectiva? Demostrar que f considerada como aplicación de T en T_+ es biyectiva.

c) $B = f(A)$ es tal que $0 < \alpha < 2$, $0 < \beta < 2$, demostrar que si se escribe

$$A_n = \sum_{q=1}^{n} \frac{(-1)^{q-1}}{q} (B - I_2)^q = \begin{pmatrix} a_n & c_n \\ 0 & b_n \end{pmatrix}$$

a_n, b_n, c_n tienen, cuando n aumenta indefinidamente, por límites respectivos a, b, c.

Concours de l'École Normale Supérieure (extracto) y M.G.P. (extracto).

209. Sea A una matriz de $M_n(\mathbf{R})$, se designa por $k_{ij}(A)$ su coeficiente situado en la i-ésima fila y la j-ésima columna. Se pone, (i, j) describiendo $[1, n] \times [1, n]$

$$N(A) = n \sup |k_{ij}(A)|$$

a) Demostrar que

$$N(A) = 0 \Rightarrow A = 0$$
$$N(\lambda A) = |\lambda| N(A) \qquad (\lambda \in \mathbf{R})$$
$$N(A \times B) \leq N(A) + N(B)$$

deducir que N(A) es una norma sobre el espacio vectorial $M_n(\mathbf{R})$. Se escribirá

$$N(A) = \|A\|$$

b) Demostrar que

$$\|AB\| \leq \|A\| \|B\|$$
$$\|(A + B)^r - A^r\| \leq [\|A\| + \|B\|]^r - \|A\|^r \qquad (r \in \mathbf{N})$$

Concours de l'École Normale Supérieure (extracto).

210. Los datos son los mismos que en el ejercicio precedente, se escribe

$$B_n = \sum_{p=0}^{n} \frac{1}{p!} A^p.$$

a) Demostrar que las propiedades precedentes son equivalentes:

α) existe una matriz B de $M_n(\mathbf{R})$, tal que

$$\lim_{n \to +\infty} \|B - B_n\| = 0$$

β) para todo par (i, j) de $[1, n] \times [1, n]$, $k_{ij}(B_n)$ tiene límite cuando $n \to +\infty$. Se pone $B = e^A$ (se introducirá $a = \sup |k_{ij}(A)|$).

b) Demostrar que

$$\|e^{A+B} - e^A\| \leq e^{\|A\|}(e^{\|B\|} - 1).$$

c) Calcular e^A para $A = \begin{pmatrix} x & z \\ 0 & y \end{pmatrix}$ (V. ej. 208).

Concours de l'École Normale Supérieure (extracto).

211. Siendo A una matriz cuadrada real, calcular e^A (V. ej. 210) en los casos siguientes:

a) $A = tI_n \quad (t \in \mathbf{R})$

b) $n = 2 \quad A = \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 0 \\ -t & 0 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & t \\ -t & 0 \end{pmatrix} \quad (t \in \mathbf{R})$

c) $n = 3 \quad A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ (V. ej. 194).

212. La matriz $A = \begin{pmatrix} x & z \\ 0 & y \end{pmatrix}$ de $M_2(\mathbf{R})$ tomada *fija*, se calcula $B = e^A$ (V. ej. 210 y 211) y se pone

$$B(t) = e^{tA} \qquad (t \in \mathbf{R}).$$

a) Demostrar que

$$B(t + t') = B(t)B(t') \qquad (t, t' \in \mathbf{R}).$$

b) Demostrar que

$$\frac{dB(t)}{dt} = AB(t) = B(t)A.$$

(Ver § 156, ej. 4.) M.G.P. (extracto).

213*. Tomando de nuevo las notaciones del ejercicio 154 (fin del capítulo 7), supondremos que los espacios vectoriales E, F. G son dimensiones finitas sobre **K**. Designando respectivamente por A_i, B_j, C_k $(1 \leq i \leq m,\ 1 \leq j \leq n,\ 1 \leq k \leq p)$ las bases de E_i, F_j, G_k tomemos por bases respectivas de E, F, G

$$A = \bigcup_{1 \leq i \leq m} A_i, \qquad B = \bigcup_{1 \leq j \leq n} B_j, \qquad C = \bigcup_{1 \leq k \leq p} C_k.$$

a) Demostrar que $M = (f, A, B)$ es un cuadro rectangular de matrices M_i^j (es decir, una «matriz de matrices») y que

$$M_i^j = M(f_i^j, A_i, B_j).$$

Determinar el número de columnas y el número de filas de M_i^j en función de las dimensiones de E_i y de F_j

b) Se escribe ($h = g \circ f$, V. ej. 154)

$$N = M(g, B, C) \qquad N_j^k = M(g_j^k, B_j, C_k)$$
$$P = M(h, A, C) \qquad P_i^k = M(h_i^k, A_i, C_k)$$

demostrar que

$$P = MN = (P_i^k) \qquad \text{con} \qquad P_i^k = \sum_{j=1}^{n} N_j^k M_i^j$$

se dice que se ha efectuado el producto de dos matrices N y M por «*bloques*».

c) Desarrollar un ejemplo en el que $m = n = 2$ y en el que las ocho matrices M_i^j, N_j^k son de tipo (2,2).

214. Sea $\mathbf{K}$ un cuerpo conmutativo, p y q dos elementos *fijos* de $\mathbf{K}$ se considera las matrices de $M_2(\mathbf{K})$.

$$A(x, y) = \begin{pmatrix} x & py \\ y & x \end{pmatrix} \qquad B(z, t) = \begin{pmatrix} z & -pt \\ t & -z \end{pmatrix}$$

y la matriz de $M_4(\mathbf{K})$ siguiente:

$$M(x, y, z, t) = \begin{pmatrix} A(x, y) & qB(z, t) \\ B(z, t) & A(x, y) \end{pmatrix}$$

Efectuar el producto por bloques (V. ej. 213) de dos matrices $M(x, y, z, t)$ $M(x', y', z', t')$ y demostrar que existe x'', y'', z'', t'' de $\mathbf{K}$ tales que

$$M(x, y, z, t)M(x', y', z', t') = M(x'', y'', z'', t'').$$

215. Si a, b, c, d son los elementos de $\mathbf{Z}$, ¿en qué condición la matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es inversible en $M_2(\mathbf{Z})$? ¿en $M_2(\mathbf{Q})$?

216. Todas las matrices consideradas pertenecen a $M_2(\mathbf{Z})$. Se considera la matriz

$$A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}.$$

a) Demostrar que A es inversible en $M_2(\mathbf{Z})$ y calcular A^{-1}.

b) Buscar las matrices X e Y de $M_2(\mathbf{Z})$ tales que

$$XA = \begin{pmatrix} -1 & 3 \\ 7 & 4 \end{pmatrix}, \qquad AY = \begin{pmatrix} 3 & -5 \\ 2 & 6 \end{pmatrix}.$$

217. a) Si A es un anillo conmutativo unitario, ¿en qué condición una matriz triangular de $M_n(A)$ es inversible en $M_n(A)$?

b) Invertir en $M(\mathbf{Z})$ las matrices ($n = 2, 3, \ldots$) ($a \in \mathbf{Z}$)

$$A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{etc.}$$

$$B_2 = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 1 & a & 0 \\ 0 & 1 & a \\ 0 & 0 & 1 \end{pmatrix} \qquad A_4 = \begin{pmatrix} 1 & a & 0 & 0 \\ 0 & 1 & a & 0 \\ 0 & 0 & 1 & a \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{etc.}$$

(se podrá razonar por inducción).

c) Demostrar que el resultado del ejercicio 193 subsiste si K es un anillo conmutativo unitario. Determinar de nuevo las inversas de las matrices A_k consideradas anteriormente escribiendo $A_k = I_k + T_k$, igualmente para B_k (observar que $(T_k)^k = 0$).

218. Si A es el anillo $\mathbf{Z}[i]$ (enteros de GAUSS, V. cap. 5, ej. 96), demostrar que la matriz

$$M = \begin{pmatrix} i & 1-i & 2+i \\ 0 & -1 & 3-i \\ 0 & 0 & -i \end{pmatrix}$$

es inversible en $M_3(A)$, calcular M^{-1}.

219. Intervenir en $M_n(\mathbf{K})$ la matriz tal que a_i^j sea igual a 0 para $i = j$ y a 1 para $i \neq j$ (razonar por recurrencia).

220. Determinar en $M_2(\mathbf{R})$ todas las matrices X que verifican $X^2 = D$, siendo D una matriz diagonal. Discutir.

221. En E de dimensión 3 sobre **K** encontrar la matriz de paso P de la base (a_i) a la base (a'_i) definida por

$$a'_1 = a_1, \qquad a'_2 = a_1 + a_2, \qquad a'_3 = a_1 + a_2 + a_3.$$

Calcular P^{-1}. Generalizar a E de dimensión n sobre **K**.

222. En $\mathbf{R}^3$ hallar la matriz de paso de la base canónica (e_i) a la base

$$e'_1 = (0, 1, 1), \qquad e'_2 = (1, 0, 1), \qquad e'_3 = (1, 1, 0).$$

Calcular P^{-1}.

223. En el espacio vectorial $\mathbf{R}^3$, se considera la aplicación lineal f cuya matriz respecto a la base canónica (e_1, e_2, e_3) es

$$\begin{pmatrix} 0 & 1 & -\operatorname{sen}\theta \\ -1 & 0 & \cos\theta \\ -\operatorname{sen}\theta & \cos\theta & 0 \end{pmatrix}$$

donde θ es un número real dado.

a) Demostrar que $f^3 = 0$.

b) A todo real t, se asocia la aplicación lineal

$$g_t = u + tf + \frac{1}{2}t^2f^2$$

donde u es la aplicación idéntica. Demostrar que el conjunto G descrito por g_t cuando t describe **R** es un grupo abeliano para la composición de las aplicaciones.

c) Se pone $e'_1 = e_1 \cos\theta + e_2 \operatorname{sen}\theta$, $e'_2 = f(e_1)$, $e'_3 = f(e_2)$, demostrar que (e'_1, e'_2, e_3) es una base de $\mathbf{R}^3$. Determinar la matriz de f en esta base.

M.G.P. (extracto).

224. Se considera la matriz del ejercicio 2 del párrafo 156 como la matriz de un endomorfismo f de **C** relativamente a la base canónica (e_1, e_2, e_3); determinar la matriz de f respecto a la base

$$e'_1 = e_1 + e_2 + e_3, \qquad e'_2 = e_2, \qquad e'_3 = e_3.$$

225. Hallar el rango de las matrices de elementos en **Q**

$$\text{a)} \begin{pmatrix} 2 & -3 & -4 \\ 3 & 1 & 5 \\ -1 & 0 & -1 \\ 0 & 2 & 4 \end{pmatrix} \qquad \text{b)} \begin{pmatrix} 1 & 7 & 5 & 3 & -2 \\ 0 & 4 & 2 & 2 & 0 \\ 2 & -2 & 4 & 0 & 1 \\ 3 & -1 & 7 & 1 & 3 \end{pmatrix}$$

226. Hallar el rango de la matriz $(p, q, r \in \mathbf{R})$

$$\begin{pmatrix} 0 & r & -q \\ -r & 0 & p \\ q & -p & 0 \end{pmatrix}.$$

227. Hallar el rango de las matrices de elementos en **C**

$$\text{a)} \begin{pmatrix} 1 & -1 & 3 \\ -1 & i & -1-2i \\ i & 1 & i-2 \end{pmatrix} \qquad \text{b)} \begin{pmatrix} 1 & -i & -i & 1 \\ i & 1 & 1 & i \\ 1 & i & 3i & 3 \end{pmatrix}.$$

228. Hallar el rango de las matrices $M_n(\mathbf{R})$

$$A_3 = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \qquad A_4 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Generalizar.

229. Hallar el rango de las matrices $M_n(\mathbf{R})$

$$A_3 = \begin{pmatrix} a & 0 & b \\ b & a & 0 \\ 0 & b & a \end{pmatrix} \qquad A_4 = \begin{pmatrix} a & 0 & 0 & b \\ b & a & 0 & 0 \\ 0 & b & a & 0 \\ 0 & 0 & b & a \end{pmatrix}.$$

Generalizar.

230. Sea $A = M(f, (a_i), (b_i))$ una matriz de tipo (m, n) sobre $\mathbf{K}$, demostrar que las diferentes aplicaciones siguientes (*«manipulación sobre las columnas»*) que asocia a A una matriz A′ del mismo tipo, son tales que $A' = AP$, donde P es una matriz cuadrada inversible de orden m que determinará en cada caso.

Se pone $A = (\alpha_i^j)$, $A' = (\alpha'^j_i)$.

a) $\alpha'^j_i = \alpha^j_{p(i)}$, p es una permutación de $[1, m]$ (P es la matriz de paso de la base (a_i) a la base $(a_{p(i)})$, es una matriz de permutación; ver ej. 204).

b) $\alpha'^j_i = \lambda^i_i \alpha^j_i$, (λ^i_i) es una familia de m escalares no nulos (P es una matriz diagonal, de elementos diagonales λ^i_i).

c) Para i *fijo*: $\alpha'^j = \alpha^j + \alpha^{j'}$ $(j' \neq j)$, las columnas $i' \neq i$ no cambian (se tiene $P = I_m + E^i_{i'}$, $E^i_{i'}$, es un elemento de la base canónica de $M_n(\mathbf{K})$, aquella en que todos los elementos son nulos salvo $\varepsilon^{i'}_i = 1$).

Deducir de ello que estas «manipulaciones sobre las columnas» dejan invariable el rango de la matriz «manipulada».

231. Los datos son los mismos que en el ejercicio anterior, demostrar que las diferentes aplicaciones siguientes (*«manipulación sobre las filas»*) que asocian a A una matriz del mismo tipo A″, son tales que $A'' = QA$, donde Q es una matriz inversible de orden n. Se pone $A'' = (\alpha''^j_i)$. Se determinará Q en cada caso.

a) $\alpha''^j_i = \alpha^{q(i)}_j$, q es una permutación de $[1, n]$.

b) $\alpha''^j_i = \mu^j_j \alpha^j_{i'}$, (μ^j_j) es una familia de n escalares no nulos.

c) Para j *fijo*: $\alpha''^j_i = \alpha^j_i + \alpha^j_{i'}$, $(i' \neq i)$ las filas $j' \neq j$ no varían.

Deducir que estas «manipulaciones sobre las líneas» dejan invariable el rango de la matriz «manipulada».

9

DETERMINANTES

Salvo indicación contraria (§ 171) los elementos de las matrices y de los determinantes empleados en este capítulo están tomados en un cuerpo conmutativo.

I. Aplicaciones y formas multilineales alternadas

163. Aplicaciones de $E_1 \times E_2 \times \ldots \times E_n$ en F

a) Aplicaciones parciales

Dados n conjuntos $E_1, E_2, \ldots, E_n$ (E_i descrito por x_i) y un $n+1$-ésimo conjunto F, generalizando lo que hemos dicho en el § 12, *d*) podemos definir una aplicación f de $E = E_1 \times E_2 \times \ldots \times E_n$ en F se escribirá

$$(x_1, x_2, \ldots, x_n) \rightarrow f(x_1, x_2, \ldots, x_n)$$

se dice también que f es una *función de las n variables* $x_1, x_2, \ldots, x_n$.

Por ejemplo, la aplicación de $E_1 \times E_2 \times \ldots \times E_n$ en E_i definida por

$$(x_1, x_2, \ldots, x_n) \rightarrow x_i$$

se llama la i-ésima *función coordenada* y se designa pr_i, se tiene luego

$$pr_i\,(x_1, x_2, \ldots, x_n) = x_i.$$

Dada una aplicación f de $E_1 \times E_2 \times \ldots \times E_n$ en F, se puede definir una y una sola aplicación g de E_i en F para cada valor del $(n-1)$-étuple $(x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ para la igualdad

$$x_i \rightarrow g(x_i) = f(x_1, x_2, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)$$

g se llama la *aplicación parcial* de E_i en F asociada a f relativamente a los valores $x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$ atribuidas a las otras $n-1$ variables; se la puede representar $f(x_1, x_2, \ldots, x_{i-1}, \cdot, x_{i+1}, \ldots, x_n)$ (ver § 12, *d*).

b) Aplicaciones simétricas

Si $E_1 = \dots = E_n = E$ se dice que f, aplicación de E^n en F es *simétrica* si para todo elemento $(x_1, x_2, \dots, x_n)$ de E^n y para toda permutación p perteneciente a $\mathfrak{S}_n$ (§ 85)

$$f(x_{p(1)}, \dots, x_{p(i)}, \dots, x_{p(n)}) = f(x_1, \dots, x_i, \dots, x_n).$$

Si se tiene la igualdad precedente únicamente para la transposición que cambia i y j $(i \neq j)$ se dirá que f es *simétrica relativamente a las variables* x_i y x_j.

c) Aplicaciones antisimétricas

Consideremos en fin una aplicación f de E^n en un *grupo*, representado aditivamente, F; se dirá que f es *antisimétrica relativamente a las variables* x_i y x_j $(i \neq j)$ si para todo elemento $(x_1, x_2, \dots, x_n)$ de E^n se tiene

$$f(x_{t(1)}, \dots, x_{t(i)}, \dots, x_{t(n)}) = -f(x_1, x_2, \dots, x_n)$$

donde t es la transposición que cambia i y j.

Si f es antisimétrica relativamente a toda pareja (x_i, x_j) $(i \neq j)$ de las n variables x_k $(1 \leq k \leq n)$ se dice que f es antisimétrica. Si p es una permutación cualquiera perteneciente a $\mathfrak{S}_n$ de signatura $\varepsilon(p)$ (§ 87), p es descomponible en producto de un número par de transposición si $\varepsilon(p) = +1$ e impar si $\varepsilon(p) = -1$, luego si f es antisimétrica

$$f(x_{p(1)}, \dots, x_{p(i)}, \dots, x_{p(n)}) = \varepsilon(p) f(x_1, \dots, x_1, \dots, x_n)$$

recíprocamente si esta igualdad se verifica para todas las transposiciones de $\mathfrak{S}_n$, se ve que f es antisimétrica respecto a toda pareja (x_i, x_j) $(i \neq j)$ de las n variables x_k; esto justifica la definición siguiente:

DEFINICIÓN 1. — *Una aplicación f de E^n en un grupo* F *(representado aditivamente) es* antisimétrico *si para toda permutación p perteneciente a* $\mathfrak{S}_n$

$$f(x_{p(1)}, \dots, x_{p(i)}, \dots, x_{p(n)}) = \varepsilon(p) f(x_1, \dots, x_i, \dots, x_n).$$

164. Aplicaciones y formas multilineales

DEFINICIÓN 2. — *Dados $n + 1$ espacios vectoriales* $E_1, E_2, \dots, E_n$, F *sobre el mismo cuerpo conmutativo* K *se dice que una aplicación f de* $E_1 \times E_2 \times \dots \times E_n$ *en* F *es una aplicación n*-lineal *si cada aplicación parcial de* E_i *en* F *asociada a f es una forma lineal.*

Si F = K *se dice que f es una* forma n-lineal.

Las aplicaciones y formas n-lineales $(n \geq 2)$ *se llaman aplicaciones y formas* multilineales.

Si $n = 2$ se dice que la aplicación o la forma es bilineal.

Si $n = 3$ se dice que la aplicación o la forma es trilineal.

Si x_i e y_i describen E_i y λ el cuerpo K se tendrá, en consecuencia, para todo i de $[1, n]$

$$f(x_1, \dots, x_{i-1}, x_i + y_i, x_{i+1}, \dots, x_n) = f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) + f(x_1, \dots, x_{i-1}, y_i, x_i, \dots, x_n)$$

$$f(x_1, \dots, x_{i-1}, \lambda x_i, x_{i+1}, \dots, x_n) = \lambda f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n).$$

Para $n = 2$ y $F = K$ se encuentran las propiedades de definición ya dadas para una forma bilineal (§ 149). Se tiene naturalmente

$$f(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n) = 0.$$

EJEMPLOS

1. La forma bilineal canónica (§ 149)

$$(x, x^*) \to \langle x, x^* \rangle$$

es una forma bilineal, aplicación bilineal de $E \times E^*$ en K.

2. Si $f^1, f^2, \dots, f^p$ son p formas lineales sobre E, espacio vectorial sobre K, la aplicación φ de E^p en K definida por

$$(x_1, \dots, x_i, \dots, x_p) \to \varphi(x_1, x_2 \dots, x_p) = f^1(x_1) \dots f^i(x_i) \dots f^p(x_p)$$

es una forma p-lineal.

Si f y g son dos aplicaciones (resp. formas) n-lineales definidas sobre $E_1 \times E_2 \times \dots \times E_n$ y con valores en F (resp. en K), $f + g$ y λf definidas, como de costumbre, por

$$(f + g)(x_1, \dots, x_i, \dots, x_n) = f(x_1, \dots, x_i, \dots, x_n) + g(x_1, \dots, x_i, \dots, x_n)$$
$$(\lambda f)(x_1, \dots, x_i, \dots, x_n) = \lambda f(x_1, \dots, x_i, \dots, x_n)$$

son aplicaciones (resp. formas) n-lineales; estas dos operaciones $f + g$ y λf verifican los axiomas de la estructura de espacio vectorial sobre K, de donde:

TEOREMA 1. — *Si* $E_1, E_2, \dots, E_n$, F *son* $n + 1$ *espacios vectoriales sobre* K, *las aplicaciones (resp. formas)* n-lineales *definidas sobre* $E_1 \times E_2 \times \dots \times E_n$ *y con valor en* F *(resp. en* K) *describen un espacio vectorial representado por* $\mathfrak{L}_n(E_1, E_2, \dots, E_n; F)$ *(resp.* $\mathfrak{L}_n(E_1, E_2, \dots, E_n; K)$) y $\mathfrak{L}_n(E; F)$ *(resp.* $\mathfrak{L}_n(E; K)$) *si*

$$E_1 = E_2 = \dots = E_n = E.$$

EJERCICIOS

1. Demostrar que el espacio vectorial $\mathfrak{L}(E, F; G)$ es isomorfo al espacio vectorial $\mathfrak{L}(E; \mathfrak{L}(F; G))$ y al $\mathfrak{L}(F; \mathfrak{L}(E; G))$.

2. Deducir del ejercicio precedente que $\mathfrak{L}(E; F; K)$ es isomorfo a $\mathfrak{L}(E; F^*)$ y a $\mathfrak{L}(F; E^*)$.

OBSERVACION

No se confundirá una aplicación n-lineal de $E_1 \times E_2 \times \dots \times E_n$ en F y una aplicación lineal del espacio vectorial $E_1 \times E_2 \times \dots \times E_n$ en F, es decir, un elemento de $\mathfrak{L}(E_1, E_2, \dots, E_n; F)$ y un elemento de $\mathfrak{L}(E_1 \times E_2 \times \dots \times E_n; F)$.

Sea, por ejemplo, una aplicación bilineal f de $E \times F$ en G; se tendrá en particular con las notaciones siguientes

$$f(x_1 + x_2, y_1 + y_2) = f(x_1, y_1) + f(x_1, y_2) + f(x_2, y_1) + f(x_2, y_2)$$
$$f(\lambda(x, y)) = f(\lambda x, \lambda y) = \lambda^2 f(x, y).$$

Por el contrario, si g es una aplicación lineal de $E \times F$ en G se tendrá

$$g(x_1 + x_2, y_1 + y_2) = g((x_1, y_1) + (x_2, y_2)) = g(x_1, y_1) + g(x_2, y_2)$$
$$g(\lambda(x, y)) = \lambda g(x, y).$$

EJERCICIO

3. Siendo f una aplicación n-lineal calcular

$$f(x_1 + y_1, \dots x_i, + y_i, \dots, x_n + y_n), \; f(\lambda x_1, \dots, \lambda x_i, \dots, \lambda x_n).$$

165. Aplicaciones y formas n-lineales alternadas

a) DEFINICIÓN 3. — *Una aplicación (resp. una forma)* n-lineal f *definida sobre* E^n *y con valores en* F (*resp.* K) *es* alternada *si es nula para todo elemento* $x = (x_1, x_2, \dots, x_n)$ *teniendo dos coordenadas iguales* (E *y* F *espacios vectoriales sobre* K).

Sea $i \neq j$, se tendrá para toda aplicación o forma multilineal alternada (suponiendo $i < j$)

$$f(x_1, \dots, x_{i-1}, x_i + x_j, x_{i+1}, \dots, x_{j-1}, x_i + x_j, x_{j+1}, \dots, x_n) = 0$$

pues los valores de la i-ésima y de la j-ésima componente son los mismos: $x_i + x_j$; de donde teniendo en cuenta de la n-linealidad

$$f(\dots, x_i, \dots, x_i, \dots) + f(\dots, x_i, \dots, x_j, \dots)$$
$$+ f(\dots, x_j, \dots, x_i, \dots) + f(\dots, x_j, \dots, x_j, \dots) = 0$$

las variables no escritas tienen los mismos valores en los cuatro términos; como f es alternada el primero y último términos son nulos, de lo que resulta

$$f(\dots, x_j, \dots, x_i, \dots) = -f(\dots, x_i, \dots, x_j, \dots)$$

de donde:

TEOREMA 2. — *Toda aplicación* (*resp. forma*) n-lineal *alternada definida sobre* E^n *con valores en* F (*resp.* K) *es antisimétrica* (*donde* E *y* F *son espacios vectoriales sobre* K); *es decir,*

$$f(x_{p(1)}, \dots, x_{p(i)}, \dots, x_{p(n)}) = \varepsilon(p) f(x_1, \dots, x_i, \dots, x_n)$$

siendo $\varepsilon(p)$ *la signatura de la permutación* p.

EJERCICIO

Demostrar, recíprocamente, que si K no es de característica 2, toda aplicación o forma n-lineal antisimétrica es alternada.

Siendo f y g dos aplicaciones (o formas) n-lineales alternadas definidas sobre E^n, es evidente que $f+g$ y λf son alternadas, de donde:

TEOREMA 3.—*Si* E *y* F *son espacios vectoriales sobre* K *el conjunto de las aplicaciones (resp. formas)* n-lineales *alternadas definidas sobre* E^n *y con valores en* F *(resp. en* K*) es un espacio vectorial sobre* K, *es un subespacio vectorial de* $\mathcal{L}_n(E; F)$ *(resp.* $\mathcal{L}_n(E; K)$).

Por otra parte, si f es una aplicación o una forma n-lineal alternada definida sobre E^n se tendrá con las notaciones precedentes (donde λx_i reemplaza x_j, $i \neq j$)

$$f(\ldots, x_i, \ldots, \lambda x_i, \ldots) = \lambda f(\ldots, x_i, \ldots, x_i, \ldots) = 0$$

igualmente

$$f\left(x_1, \ldots, x_{i-1}, \sum_{k=1}^{n} \lambda^k x_k, x_{i+1}, \ldots, x_n\right) = \sum_{k=1}^{n} \lambda^k f(x_1, \ldots, x_{i-1}, x_k, x_{i+1}, \ldots, x^n)$$
$$= \lambda^i f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n).$$

Si la familia $(x_1, x_2, \ldots, x_n)$ es ligada existirá una familia de escalares (λ^k) no todos nulos tales que $\sum_{k=1}^{n} \lambda^k x_k = 0$, si, por ejemplo, $\lambda^i \neq 0$ se tendrá

$$f\left(x_1, \ldots, x_{i-1}, \sum_{k=1}^{n} \lambda^k x_k, x_{i+1}, \ldots, x_n\right) = f(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) = 0$$
$$= \lambda^i f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)$$

luego:

TEOREMA 4.—*Siendo* f *una aplicación o una forma* n-lineal *alternada definida sobre* E^n, *si* $x_1, \ldots, x_n$ *es una familia ligada de* E, $f(x_1, \ldots, x_n) = 0$.

En este párrafo hemos dado las propiedades de las formas n-lineales alternadas definidas sobre E^n, pero no hemos demostrado que existen (aparte del caso trivial $f = 0$). Vamos a hacerlo en el párrafo siguiente, en el único caso que nos interesa en este curso: aquel en que la dimensión de E es n.

166. Formas n-lineales alternadas definidas sobre E^n (dim E $= n$)

a) Caso en que $n = 2$

Empecemos por estudiar el caso de $n = 2$; sea $\{a, b\}$ una base de E, pongamos

(1) $$x = \alpha a + \beta b, \qquad y = \alpha' a + \beta' b$$

si f es una forma bilineal alternada definida sobre E, tendremos

$$f(x, y) = f(\alpha a + \beta b, \alpha' a + \beta' b)$$
$$= \alpha\alpha' f(a, a) + \alpha\beta' f(a, b) + \alpha'\beta f(b, a) + \beta\beta' f(b, b)$$

(2) $$f(x, y) = (\alpha\beta' - \alpha'\beta) f(a, b)$$

ahora bien, la aplicación d de E^2 en K definida por

(2′) $$d(x, y) = \alpha\beta' - \alpha'\beta$$

es manifiestamente bilineal y alternada; además, $d(a, b) = 1$: existen, pues, formas bilineales alternadas definidas sobre E^2 no nulas (por ejemplo, d), y la fórmula muestra que toda forma bilineal alternada f definida sobre E^2 es tal que

(3) $$f = \lambda d \qquad \lambda \in K$$

luego *el espacio vectorial de las formas bilineales definidas sobre* E^2 *es de dimensión* 1. Recíprocamente la fórmula (3) muestra que para todo valor λ de K hay una forma bilineal alternada única tal que

$$f(a, b) = \lambda.$$

b) **Caso general**

Consideremos ahora las formas n-lineales alternadas definidas sobre E^n ($\dim E = n$), suponiendo que exista. Los razonamientos son exactamente los mismos que en el caso particular precedente ($n = 2$), pero el hecho de que la dimensión sea n complica los cálculos.

Sea $\{a_1, \ldots, a_n\}$ una base de E y f una forma n-lineal alternada definida sobre E^n; debiendo considerar un elemento cualquiera $(x_1, \ldots, x_n)$ de E^n, es decir, debiendo coger simultáneamente n vectores de E, para cada vector x_i $(1 \leq i \leq n)$

$$x_i = \sum_{j=1}^{n} \alpha_i^j a_j$$

representaremos el índice sumatorio (es decir, el número de la coordenada) por una letra relativa a cada vector (ver § 91, *e*); pondremos, pues,

(4) $$i \in [1, n] \qquad x_i = \sum_{j_i=1}^{n} \alpha_i^{j_i} a_{j_i}$$

Calculamos $f(x_1, \ldots, x_n)$ utilizando primero las *dos propiedades de n-linealidad de f*

$$f(x_1, \ldots, x_i, \ldots, x_n) = f\left(\sum_{j_1=1}^{n} \alpha_1^{j_1} a_{j_1}, \ldots, \sum_{j_i=1}^{n} \alpha_i^{j_i} a_{j_i}, \ldots, \sum_{j_n=1}^{n} \alpha_n^{j_n} a_{j_n}\right)$$
$$= \sum_{j_1=1}^{n} \cdots \sum_{j_i=1}^{n} \cdots \sum_{j_n=1}^{n} f(\alpha_1^{j_1} a_{j_1}, \ldots, \alpha_i^{j_i} a_{j_i}, \ldots, \alpha_n^{j_n} a_{j_n})$$
$$= \sum_{j_1=1}^{n} \cdots \sum_{j_i=1}^{n} \cdots \sum_{j_n=1}^{n} \alpha_1^{j_1} \ldots \alpha_i^{j_i} \ldots \alpha_n^{j_n} f(a_{j_1}, \ldots, a_{j_i}, \ldots, a_{j_n})$$
$$= \sum_{(j_1, \ldots, j_i, \ldots, j_n) \in I^n} \alpha_1^{j_1} \ldots \alpha_i^{j_i} \ldots \alpha_n^{j_n} f(a_{j_1}, \ldots, a_{j_i}, \ldots, a_{j_n})$$

en la penúltima forma $f(x_1, \ldots, x_n)$ es una Σ n-étuple, y en la última una Σ simple, pero cuyo índice sumatorio describe I^n; hay, pues, n^n términos, mas muchos de ellos son nulos: en efecto, utilicemos ahora el hecho que f es *alternada*: sea $i \neq i'$, por ejemplo, $i < i'$,

$$j_i = j_{i'} \Rightarrow f(a_{j_1}, \ldots, a_{j_i}, \ldots, a_{j_{i'}}, \ldots, a_{j_n}) = 0.$$

Si en la expresión $f(x_1, \ldots, x_n)$ suprimimos estos términos nulos, sólo quedarán los términos tales que $i \rightarrow j_i$ sea inyectiva, luego biyectiva (§ 31, corolario 3 del teorema 4); para estos términos tenemos que

$$j_i = p(i) \qquad \text{con} \qquad p \in \mathfrak{S}_n.$$

Como $\mathfrak{S}_n$ tiene $n!$ elementos (§ 85) ello nos dice que en $f(x_1, \ldots, x_n)$ sólo existen $n!$ términos, y que todos los demás son nulos, escribiremos

$$f(x_1, \ldots, x_n) = \sum_{p \in \mathfrak{S}_n} \alpha_1^{p(1)} \ldots \alpha_i^{p(i)} \ldots \alpha_n^{p(n)} f(a_{p(1)}, \ldots, a_{p(i)}, \ldots, a_{p(n)})$$

donde la notación indica que la suma se extiende a todos los términos cada uno correspondiente a una permutación p de $\mathfrak{S}_n$.

En fin, f, forma n-lineal alternada, es *antisimétrica* (§ 165, teorema 2); tenemos, pues, con $\varepsilon(p)$ la signatura de p

$$f(x_1, \ldots, x_n) = \sum_{p \in \mathfrak{S}_n} \alpha_1^{p(1)} \ldots \alpha_i^{p(i)} \ldots \alpha_n^{p(n)} \varepsilon(p) f(a_1, \ldots, a_i, \ldots, a_n)$$

de donde finalmente

$$(5) \qquad f(x_1, \ldots, x_n) = f(a_1, \ldots, a_n) \sum_{p \in \mathfrak{S}_n} \varepsilon(p)\, \alpha_1^{p(1)} \ldots \alpha_i^{p(i)} \ldots \alpha_n^{p(n)}.$$

Recíprocamente elegida la base $\{a_1, \ldots, a_n\}$, consideremos la aplicación d de E en K definida por

$$(5') \qquad d(x_1, \ldots, x_n) = \sum_{p \in \mathfrak{S}_n} \varepsilon(p)\, \alpha_1^{p(1)} \ldots \alpha_i^{p(i)} \ldots \alpha_n^{p(n)}$$

cada término de la Σ depende linealmente de cada uno de los vectores $x_1, \ldots, x_n$, luego d es una forma n-lineal; por otra parte, $d \neq 0$, pues $d(a_1, \ldots, a_n) = 1$; demostremos en fin que d es *alternada*: sea i e i' dos enteros de $[1, n]$ distintos, nos proponemos demostrar que $x_i = x_{i'}$ implica $d(x_1, \ldots, x_n) = 0$.

Sea t la transposición tal que $t(i) = i'$, se tiene $t(i') = i$ y para todo k de $[1, n]$ distinto de i e i', $t(k) = k$ (ver § 85). Designemos por H el subgrupo de $\mathcal{S}_n$ engendrado por t, tenemos $t^2 = u$ (u: identidad), luego $\mathrm{H} = \{u, t\}$. Cada clase por la izquierda según el subgrupo H tiene dos elementos y el conjunto de las clases por la izquierda es una partición de $\mathcal{S}_n$; si p es uno de los elementos de una clase por la izquierda, la otra será $p' = pt$ (§ 74) y si p describe P, p' describirá P′ con

$$\mathrm{P} \cap \mathrm{P}' = \varnothing \quad \text{y} \quad \mathrm{P} \cup \mathrm{P}' = \mathcal{S}_n$$

luego

$$d(x_1, \ldots, x_n) = \sum_{p \in \mathrm{P}} \varepsilon(p)\, \alpha_1^{p(1)} \ldots \alpha_n^{p(n)} + \sum_{p' \in \mathrm{P}'} \varepsilon(p')\, \alpha^{p'(1)} \ldots \alpha_n^{p'(n)}$$

o también, observando (§ 87) que,

$$\varepsilon(p') = \varepsilon(pt) = \varepsilon(p)\,\varepsilon(t) = -\varepsilon(p)$$

$$d(x_1, \ldots, x_n) = \sum_{p \in \mathrm{P}} \varepsilon(p)\, [\alpha_1^{p(1)} \ldots \alpha_n^{p(n)} - \alpha_1^{(pt)(1)} \ldots \alpha_n^{(pt)(n)}]$$

si $k \neq i$ y $k \neq i'$

$$(pt)(k) = p(k), \qquad \text{pues} \qquad t(k) = k,$$

si $k = i$

$$(pt)(i) = p(i'), \qquad \text{pues} \qquad t(i) = i',$$

si $k = i'$

$$(pt)(i') = p(i), \qquad \text{pues} \qquad t(i') = i.$$

En consecuencia, $x_i = x_{i'}$ contiene cualquiera que sea k de $[1, n]$

$$\alpha_k^{pt(k)} = \alpha_k^{p(k)}, \quad \alpha_i^{pt(i)} = \alpha_i^{p(i')} = \alpha_{i'}^{p(i')}, \quad \alpha_{i'}^{pt(i')} = \alpha_{i'}^{p(i)} = \alpha_i^{p(i)}$$

y todos los términos del corchete de la última Σ son nulos: *d es, por tanto, una forma* n-lineal *alternada definida sobre* E^n *y es no nula, ya que* $d(a_1, \ldots, a_n) = 1$.

Según (5) y (6) se tiene, poniendo $f(a_1, \ldots, a_n) = \lambda$

$$f(x_1, \ldots, x_n) = \lambda d(x_1, \ldots, x_n) \qquad \lambda \in \mathrm{K}$$

para todo elemento $(x_1, \ldots, x_n)$ de E^n; en consecuencia,

$$f = \lambda d. \tag{6}$$

Luego d engendra el espacio vectorial de las formas n-lineales alternadas definidas sobre E^n, que es, pues, de dimensión 1. Por otro lado, como $d(a_1, \ldots, a_n) = 1$, la fórmula (6) nos dice que para todo λ de K hay una forma n-lineal alternada única tal que $f(a_1, \ldots, a_n) = \lambda$; de donde:

TEOREMA 5. — *Siendo* E *un espacio vectorial de dimensión* n *sobre* K, *el espacio vectorial de las formas* n-lineales *alternadas definidas sobre* E^n *es de dimensión* 1. *Para todo* λ *de* K *existe una forma* n-lineal *alternada única* f *tal que*

$$f(a_1, \ldots, a_n) = \lambda$$

siendo $\{a_1, \ldots, a_n\}$ *una base de* E.

La fórmula (5) demuestra que si para una base (a_i) $f(a_1, \ldots, a_n) = 0$, se tendrá $f(x_1, \ldots, x_n) = 0$ para todo elemento de E^n, luego $f = 0$. De ello resulta que si (a_i) y (b_i) son dos bases cualesquiera de E la condición $f(b_1, \ldots, b_n) \neq 0$ implica $f(a_1, \ldots, a_n) \neq 0$, de donde:

COROLARIO 1. — *Si una forma* n-lineal *alternada definida sobre* E^n (dim E $= n$), *toma un valor no nulo para una base, toma también un valor no nulo para toda base de* E.

Se podría igualmente decir que $f = 0$ si y sólo si f toma el valor cero para una base particular. Luego si $f \neq 0$ y $f(x_1, \ldots, x_n) = 0$, entonces $\{x_1, \ldots, x_n\}$ es una familia ligada, sino $\{x_1, \ldots, x_n\}$ sería una base y f sería nula, de donde:

COROLARIO 2. — *Si* f *es un forma* n-lineal *alternada no nula definida sobre* E^n (dim E $= n$), $f(x_1, \ldots, x_n) = 0$ *si y sólo si* $\{x_1, \ldots, x_n\}$ *es una parte ligada de* E.

EJERCICIOS

1. Sea f una forma bilineal alternada definida sobre E^2, con E de dimensión n sobre K, expresado en una base (a_i). Se tiene

$$x = \sum_{i=1}^{n} \alpha^i a_i, \qquad y = \sum_{j=1}^{n} \beta^j a_j, \qquad \varphi_{ij} = f(a_i, a_j),$$

demostrar que

$$f(x, y) = \sum_{1 \leq i < j \leq n} (\alpha^i \beta^j - \alpha^j \beta^i)\varphi_{ij},$$

recíprocamente, demostrar que la fórmula precedente, cuando se da arbitrariamente los escalares φ_{ij} $(1 \leq i < j \leq n)$ define una forma bilineal alternada definida sobre E^2. Deducir la dimensión del espacio vectorial de las formas bilineales alternadas definidas sobre E^2. Demostrar que para $E = K^3$ este espacio es isomorfo a E.

2. Siendo E de dimensión n sobre K, encontrar la dimensión del espacio vectorial de las formas m-lineales alternadas definidas sobre E^m $(m < n)$.

3. Sea f una forma n-lineal definida sobre E^n demostrar que la aplicación g de E^n en K definida por

$$g(x_1, \ldots, x_i, \ldots, x_n) = \sum_{p \in S_n} \varepsilon(p) f(x_{p(1)}, \ldots, x_{p(i)}, \ldots, x_{p(n)})$$

es una forma n-lineal alternada definida sobre E^n. (Se dice que la forma n-lineal alternada g *se ha obtenido por antisimetrización* de la forma n-lineal f.)

II. Determinantes

167. Definición y cálculo de un determinante

a) El teorema 5 del párrafo precedente nos permite enunciar:

Definición. — *Siendo* E *un espacio vectorial de dimensión* n *sobre* K *referido a una base* (a_i) $(1 \leq i \leq n)$, *existe una y sólo una forma* n-lineal *alternada definida sobre* E^n *que toma el valor* 1 *para* $(a_1, \ldots, a_n)$. *Su valor para* $(x_1, \ldots, x_n)$ *se llama* determinante *de* $(x_1, \ldots, x_n)$ *respecto a la base* $(a_1, \ldots, a_n)$, *se le representa por*

$$\det_{(a_i)} (x_1, \ldots, x_n)$$

o $\det (x_1, \ldots, x_n)$ *si no da lugar a confusión.*

Por abuso de lenguaje se llama también determinante (relativamente a la base (a_i)) la aplicación $\det_{(a_i)}$ de E^n en K así definida, que no es otra que la aplicación n-lineal alternada d definida en el párrafo precedente.

Si se pone

$$1 \leq i \leq n \qquad x_i = \sum_{j=1}^{n} \alpha_i^j a_j$$

tenemos (§ 166)

$$(1) \qquad \det_{(a_i)}(x_1, \ldots, x_i, \ldots, x_n) = \sum_{p \in S_n} \varepsilon(p) \alpha_1^{p(1)} \ldots \alpha_i^{p(i)} \ldots \alpha_n^{p(n)}$$

lo que se escribe en la forma

$$(1') \qquad \det_{(a_i)}(x_1, \ldots, x_i, \ldots, x_n) = \begin{vmatrix} \alpha_1^1 & \ldots & \alpha_i^1 & \ldots & \alpha_n^1 \\ \vdots & & & & \vdots \\ \alpha_1^j & \ldots & \alpha_i^j & \ldots & \alpha_n^j \\ \vdots & & & & \vdots \\ \alpha_1^n & \ldots & \alpha_i^n & \ldots & \alpha_n^n \end{vmatrix}.$$

b) **Determinante de una matriz cuadrada A de orden n sobre K**

Los n^2 escalares α_i^j definen n vectores (x_i) de un espacio vectorial E referido a la base (a_i) mediante las fórmulas $x_i = \sum_{j=1}^{n} \alpha_i^j a_j$, se dirá, pues, que el determinante de los n vectores (x_i) respecto a la base (a_i) es el *determinante de la matriz* A y se escribirá

$$\det A = \begin{vmatrix} \alpha_1^1 & \ldots & \alpha_n^1 \\ \vdots & & \vdots \\ \alpha_1^n & \ldots & \alpha_n^n \end{vmatrix} = \det (\alpha^j)$$

en consecuencia: det A *es el determinante de los n vectores columnas de* A, *vectores de* K^n *expresado en su base canónica.* Por otra parte, si se considera A como matriz de un endomorfismo f se tiene

$$A = M(f, (a_i)) = (\alpha_i^j)$$

con

$$f(a_i) = \sum_{j=1}^{n} \alpha_i^j a_j$$

de donde

(2) $$\det A = \det M(f, (a_i)) = \det_{(a_i)}(f(a_1), \ldots, f(a_n))$$

en particular cualquiera que sea la base (a_i).

(3) $$\det I_n = \det M(\mathrm{id}_E, (a_i)) = \det_{(a_i)}(a_1, \ldots, a_n) = 1.$$

Si A es de orden n, se dice igualmente que det A es de *orden* n, como la matriz A está formado con la ayuda de n^2 *elementos* α_i^j; es un escalar que es la suma de los $n!$ *términos* $\varepsilon(p)\alpha_1^{p(1)} \ldots \alpha_n^{p(n)}$:

EJEMPLOS

1. *Determinante de orden* 2

$$\begin{vmatrix} \alpha_1^1 & \alpha_2^1 \\ \alpha_1^2 & \alpha_2^2 \end{vmatrix} = \alpha_1^1\alpha_2^2 - \alpha_1^2\alpha_2^1$$

puesto que la permutación $\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$ es par y la permutación $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ impar.

Nos encontramos naturalmente con el resultado hallado en el § 166, *a*)

$$\begin{vmatrix} \alpha & \alpha' \\ \beta & \beta' \end{vmatrix} = \alpha\beta' - \beta\alpha'$$

se obtendrá la *regla*: *Un determinante de orden 2 es igual al producto de los elementos de la diagonal principal menos el producto de los elementos de la diagonal no principal*.

2. *Determinante de orden* 3. — Hemos visto en el § 87 que entre las $3! = 6$ permutaciones de $\mathfrak{S}_3$ hay 3 que son pares

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

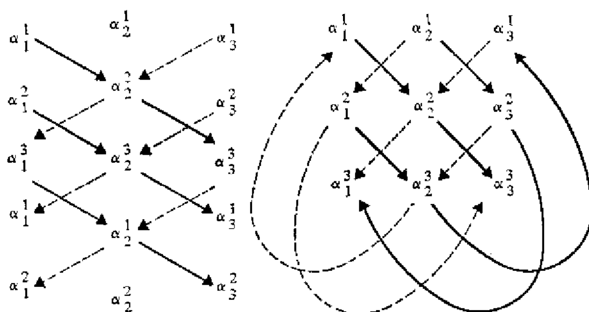
y 3 que son impares

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$

de donde

$$\begin{vmatrix} \alpha_1^1 & \alpha_2^1 & \alpha_3^1 \\ \alpha_1^2 & \alpha_2^2 & \alpha_3^2 \\ \alpha_1^3 & \alpha_2^3 & \alpha_3^3 \end{vmatrix} = \alpha_1^1\alpha_2^2\alpha_3^3 + \alpha_1^2\alpha_2^3\alpha_3^1 + \alpha_1^3\alpha_2^1\alpha_3^2 - (\alpha_1^1\alpha_2^3\alpha_3^2 + \alpha_1^3\alpha_2^2\alpha_3^1 + \alpha_1^2\alpha_2^1\alpha_3^3)$$

fórmula materializada por uno de los esquemas siguientes



los productos de los elementos unidos por una flecha continua están precedidos del signo + y los productos de los elementos unidos por una flecha de trazos están precedidos del signo — (regla de SARRUS).

Para los determinantes de orden $n > 3$ la aplicación de la fórmula general (1) es en general pesado, indicaremos en el § 170 un procedimiento que permite llevar el cálculo de un determinante de orden n al de determinantes de orden $n - 1$, de donde por inducción al cálculo de determinantes de orden 3 o 2.

EJERCICIOS

1. Calcular

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ \alpha^2 & \beta^2 & \gamma^2 \end{vmatrix} \qquad \begin{vmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & 1 & \lambda \end{vmatrix} \qquad \begin{vmatrix} \alpha & \beta & \gamma \\ \gamma & \alpha & \beta \\ \beta & \gamma & \alpha \end{vmatrix}$$

2. Demostrar que el determinante de una matriz triangular es igual al producto de los elementos diagonales de esta matriz (utilizar la fórmula (1), ver otra demostración en el § 170, *b*).

168. Primeras propiedades de los determinantes

a) Regla de dualidad

Sea

$$\det \text{A} = \sum_{p \in S_n} \varepsilon(p)\, \alpha_1^{p(1)} \dots \alpha^{p(i)} \dots \alpha_n^{p(n)}$$

Sea q una permutación cualquiera de S_n, como K es conmutativo,

$$\alpha_1^{p(1)} \dots \alpha_i^{p(i)} \dots \alpha_n^{p(n)} = \alpha_{q(1)}^{p(q(1))} \dots \alpha_{q(i)}^{p(q(i))} \dots \alpha_{q(n)}^{p(q(n))}$$

Tomemos en particular $q = p^{-1}$, se tendrá $p(p^{-1}(i)) = i$; por otra parte (§ 87),

$$\varepsilon(p) = \varepsilon(p^{-1})$$

de donde

$$\det A = \sum_{p \in \mathfrak{S}_n} \varepsilon(p^{-1})\, \alpha^1_{p^{-1}(1)} \cdots \alpha^i_{p^{-1}(i)} \cdots \alpha^n_{p^{-1}(n)} = \sum_{p \in \mathfrak{S}_n} \varepsilon(p)\, \alpha^1_{p(1)} \cdots \alpha^i_{p(i)} \cdots \alpha^n_{p(n)}$$

pues p^{-1} es, igual que p, el elemento general de $\mathfrak{S}_n$. Luego si se pone $\alpha'^j_i = \alpha^i_j$ para todo par (i, j) se obtiene

$$\det A = \det {}^tA$$

diremos que un determinante conserva el mismo valor cuando "*se cambia las columnas por las filas*". Se ve también que det A es el determinante de los *vectores filas de* A, *elementos de* $(K^n)^*$ respecto a la base dual de la base canónica de K^n.

b) Designando por (c_i) la familia de los vectores columnas de A y por (l^j) la familia de los vectores filas de A, tenemos

$$\det A = \det(c_1, \ldots, c_n) = \det(l^1, \ldots, l^n).$$

El hecho de que la aplicación $(x_1, \ldots, x_n) \to \det(x_1, \ldots, x_n)$ sea n-lineal alternada implica un cierto número de propiedades que hemos ya enunciado; para la comodidad del lector las agrupamos aquí. Según la regla de dualidad las propiedades de las columnas se aplican a las filas, a fin de no repetir dos veces los enunciados designaremos por *líneas paralelas* tanto a dos columnas como a dos líneas

1. $\det(c_1, \ldots, c_i + c'_i, \ldots, c_n) = \det(c_1, \ldots, c_i, \ldots, c_n) + \det(c_1, \ldots, c'_i, \ldots, c_n)$

$$\det(c_1, \ldots, \lambda c_i, \ldots, c_n) = \lambda \det(c_1, \ldots, c_i, \ldots, c_n)$$
$$\det(l^1, \ldots, l^j + l'^j, \ldots, l^n) = \det(l^1, \ldots, l^j, \ldots, l^n) + \det(l^1, \ldots, l'^j, \ldots, l^n)$$
$$\det(l^1, \ldots, \mu l^j, \ldots, l^n) = \mu \det(l^1, \ldots, l^j, \ldots, l^n).$$

2. Toda permutación p (elemento de $\mathfrak{S}_n$) sobre las columnas (o las filas) transforma det A en $\varepsilon(p)$ det A.

En particular toda transposición sobre las columnas (o las filas) transforma det A en — det A.

3. Un determinante es nulo (caso particular de la propiedad 5 más abajo):

— Si una columna (o una fila) es nula.

— Si dos líneas paralelas de números diferentes son iguales o más generalmente proporcionales.

4. Para toda familia (λ^k) o (μ_k) de escalares $(1 \leq k \leq n)$

$$\det\left(c_1, \ldots, c_{i-1}, \sum_{k=1}^{n} \lambda^k c_k, c_{i+1}, \ldots, c_n\right) = \lambda^i \det(c_1, \ldots, c_i, \ldots, c_n)$$

$$\det\left(l^1, \ldots, l^{j-1}, \sum_{k=1}^{n} \mu_k l^k, l^{j+1}, \ldots, l^n\right) = \mu_j \det(l^1, \ldots, l^j, \ldots, l^n)$$

en particular: si se quiere conservar el valor de un determinante reemplazando una línea por una combinación lineal de líneas paralelas *es necesario que el multiplicador de la línea reemplazada sea* 1.

5. Finalmente un determinante es nulo si y sólo si las columnas (o las filas) forman una familia ligada (corolario 2 del teorema 5 del § 166).

EJERCICIO

Si K es un cuerpo de característica diferente de 2, el determinante de toda matriz antisimétrica de orden impar es nulo.

169. Determinante del producto de dos matrices. Aplicaciones

a) Determinante del producto de dos matrices

Sea E un espacio vectorial de dimensión n sobre K referido a una base (a_i) $(1 \le i \le n)$, consideremos dos endomorfismos f y g de E y pongamos

$$A = M(f,\ (a_i)) \qquad B = M(g,\ (a_i)).$$

Consideremos la aplicación de E^n en K

$$(x_1,\ \ldots,\ x_i,\ \ldots,\ x_n) \to \det_{(a_i)} (g(x_1),\ \ldots,\ g(x_i),\ \ldots,\ g(x_n))$$

es n-lineal; por otra parte, es alternada, pues $x_i = x_{i'}$ implica $g(x_i) = g(x_{i'})$. En consecuencia, esta aplicación es una forma n-lineal alternada definida sobre E^n.

Por otra parte, podemos tomar como base del espacio vectorial de las formas n-lineales alternadas definidas sobre E^n, $d = \det_{(a_i)}$ que es no nulo, puesto que $d(a_1,\ \ldots,\ a_n) = 1$, luego (§ 166, teorema 5)

$$(1) \qquad \det_{(a_i)} [g(x_1),\ \ldots,\ g(x_i),\ \ldots,\ g(x_n)] = \lambda \det_{a_i} (x_1,\ \ldots,\ x_i,\ \ldots,\ x_n)$$

fórmula que es válida cualquiera que sea el elemento $(x_1,\ \ldots,\ x_i,\ \ldots,\ x_n)$ de E^n; dicho de otro modo, λ no depende de $(x_1,\ \ldots,\ x_i,\ \ldots,\ x_n)$, sino únicamente de g.

Tendremos tomando $x_i = a_i$ $(1 \le i \le n)$

$$(2) \quad \det_{(a_i)} [g(a_1),\ \ldots,\ g(a_i),\ \ldots,\ g(a_n)] = \lambda \det_{(a_i)} (a_1,\ \ldots,\ a_i,\ \ldots,\ a_n) = \lambda.$$

Tomemos ahora $x_i = f(a_i)$, $(1 \le i \le n)$; tendremos

$$(3) \qquad \begin{aligned} &\det_{(a_i)} [(g \circ f)(a_1),\ \ldots,\ (g \circ f)(a_i),\ \ldots,\ (g \circ f)(a_n)] \\ &\quad = \lambda \det_{(a_i)} [f(a_1),\ \ldots,\ f(a_i),\ \ldots,\ f(a_n)] \end{aligned}$$

de donde

$$(4) \qquad \begin{aligned} &\det\ [(g \circ f)(a_1),\ \ldots,\ (g \circ f)(a_i),\ \ldots,\ (g \circ f)(a_n)] \\ &\quad = \det\ [g(a_1),\ \ldots,\ g(a_n)] \det\ [f(a_1),\ \ldots,\ f(a_n)] \end{aligned}$$

de donde finalmente según la fórmula (2) del § 167, *b*)

$$\det_{(a_i)} M(g \circ f,\ (a_i)) = \det_{(a_i)} M(g,\ (a_i)) \det_{(a_i)} M(f,\ (a_i)),$$

es decir,

$$\det (BA) = \det B \det A.$$

Como K es conmutativo y como $\det {}^tA = \det A$ tenemos

$$\det (BA) = \det (AB) = \det ({}^tBA) = \det (A{}^tB) = \det (B{}^tA)$$
$$= \det ({}^tAB) = \det ({}^tB{}^tA) = \det ({}^tA{}^tB).$$

EJERCICIO

1. $A = (\alpha_i^j)$ es una matriz cuadrada de orden n tal para todo i y todo i' de $[1, n]$ se tenga

$$\sum_{j=1}^{n} \alpha_i^j \alpha_j^{i'} = \delta_i^{i'}$$

siendo $\delta_i^{i'}$ el símbolo de KRONECKER, demostrar que $\det (A^2) = 1$.

b) Caso de dos matrices inversas

Si A es inversible, existe A^{-1} tal que

$$AA^{-1} = A^{-1}A = I_n$$

de donde

$$\det A \cdot \det (A^{-1}) = 1$$

luego si A es inversible, $\det A \neq 0$ y los escalares $\det A$ y $\det (A^{-1})$ son inversos en K. Recíprocamente sea A una matriz cuadrada de orden n tal que $\det A \neq 0$, se puede considerar A como la matriz asociada a un endomorfismo f de un espacio E respecto a una base (a_i) $(1 \leq i \leq n)$, luego (§ 167, *b*), fórmula (2))

$$\det A = \det_{(a_i)} [f(a_1), \ldots, f(a_i), \ldots, f(a_n)] \neq 0$$

los n vectores $f(a_i)$ son, pues, linealmente independientes, luego $\dim f(E) = n$; f es suprayectiva y, en consecuencia, un automorfismo de E (§ 143) y $A = M(f)$ es inversible (§ 158, *c*), de donde reuniendo todos estos resultados:

TEOREMA 6. — *Si f es un endomorfismo de un espacio vectorial* E *de dimensión n sobre* K, *y (a_i) una base de* E, *las propiedades siguientes son equivalentes:*

1. *f es inyectiva.*
2. *f es suprayectiva.*
3. *f es un automorfismo de* E *(o un operador lineal regular).*
4. $A = M(f, (a_i))$ *es inversible.*
5. $\det A \neq 0$.

Hemos visto que si la familia (x_i) $(1 \leq i \leq n)$ es ligada, el determinante de estos n vectores es nulo; considerando el endomorfismo f definido por $f(a_i) = x_i$, tenemos las dos proposiciones (contrapuestas una de la otra, luego equivalentes; ver § 2)

$$\det_{(a_i)} (x_1, \ldots, x_i, \ldots, x_n) \neq 0 \Rightarrow (x_i) \text{ es una familia libre}$$
$$\det_{(a_i)} (x_1, \ldots, x_i, \ldots, x_n) = 0 \Rightarrow (x_i) \text{ es una familia ligada.}$$

Habíamos ya obtenido estos resultados (§ 168, *b*), propiedad 5).

c) Determinantes de dos matrices semejantes. Determinante de un endomorfismo

Sea A y B dos matrices cuadradas semejantes de orden n (§ 161, *c*), existe una matriz cuadrada inversible P tal que

$$B = P^{-1}AP$$

de donde

$$\det B = (\det P^{-1})(\det A)(\det P) = (\det P)^{-1}(\det P)(\det A) = \det A$$

ya que el cuerpo K es conmutativo, de donde:

TEOREMA. — *Todas las matrices cuadradas semejantes sobre un cuerpo* K *conmutativo tienen el mismo determinante.*

Sea f un endomorfismo de E, las matrices asociadas a f en dos bases cualesquiera de E son semejantes (§ 161, *c*), de donde:

COROLARIO Y DEFINICIÓN. — *Si f es un endomorfismo de un espacio vectorial* E *de dimensión n sobre* K, $\det M(f, (a_i))$, *tiene un valor independiente de la base escogida* (a_i); *se dice que este determinante es el determinante del endomorfismo f y se le representa por* $\det(f)$.

Se tiene, en consecuencia,

$$\det(g \circ f) = (\det g)(\det f).$$

EJERCICIO

2. ¿Qué se puede decir de la aplicación $f \to \det(f)$ de $\mathfrak{L}_K(E)$ en K, o de $\mathbf{GL}_n(K)$ en K?

170. Desarrollo de un determinante con relación a los elementos de una columna (o de una fila)

a) Adjuntos y menores

Designemos por $x_1, \ldots, x_i, \ldots, x_n$ los vectores columnas de la matriz de orden n, $A = (\alpha_i^j)$, la aplicación $f = \det.(x_1, \ldots, x_{i-1}, ., x_{i+1}, \ldots, x_n)$ de E en K definida por

$$x_i \to f(x_i) = \det(x_1, \ldots, x_i, \ldots, x_n) = \det(\alpha_i^j)$$

es lineal; es, pues, una forma lineal definida sobre E, y existen los escalares $\beta_1^i, \ldots, \beta_j^i, \ldots, \beta_n^i$ tales que

$$(1) \qquad \det A = \sum_{j=1}^{n} \beta_j^i \alpha_i^j = \beta_1^i \alpha_i^1 + \ldots + \beta_j^i \alpha_i^j + \ldots + \beta_n^i \alpha_i^n$$

estos escalares β_j^i $(1 \leq j \leq n)$ sólo dependen de la forma f, es decir, de $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n$; no dependen de x_i y, en consecuencia, tampoco de

$$\alpha_i^1, \ldots, \alpha_i^j, \ldots, \alpha_i^n.$$

Consideremos en det A, la suma de los términos conteniendo α_i^j, es decir, los términos de la forma $\varepsilon(p)\alpha_1^{p(1)} \ldots \alpha_i^{p(i)} \ldots \alpha_n^{p(n)}$ con $p(i) = j$; si fijamos j su suma es

$$\beta_j^i \alpha_i^j = \sum_{p \in S_n, p(i)=j} \varepsilon(p)\alpha_1^{p(1)} \ldots \alpha_i^{p(i)} \ldots \alpha_n^{p(n)}$$

puesto que β_j^i no depende de x_i, la fórmula precedente es verdadera cualquiera que sea x_i podemos suponer $\alpha_i^j \neq 0$, de donde

$$\beta_j^i = \sum_{p \in S_n, p(i)=j} \varepsilon(p)\alpha_1^{p(1)} \ldots \alpha_{i-1}^{p(i-1)}\alpha_{i+1}^{(p+1)} \ldots \alpha_n^{p(n)}$$

$p(i) = j$, situado debajo de Σ, recuerda que, siendo p una permutación de $[1, n]$, cada uno de los índices superiores $p(1), \ldots, p(i-1), p(i+1), \ldots p(n)$ pertenecen a $[1, n] - \{j\}$.

Según la regla de dualidad se puede también desarrollar det A respecto a los elementos de la fila j tendremos

$$(1') \qquad \det A = \sum_{i=1}^{n} \gamma_j^i \alpha_i^j = \gamma_j^i \alpha_1^j + \ldots + \gamma_j^i \alpha_i^j + \ldots + \gamma_j \alpha_n^j.$$

La suma de los términos del determinante conteniendo α_i^j es evidentemente la misma en los dos desarrollos (1) y (1'), luego $\beta_j^i = \gamma_j^i$; este coeficiente β_j^i de α_i^j, en el desarrollo de det A respecto a los elementos de la columna i o de la fila j, se llama el *adjunto* de α_i^j (observar el lugar de los índices en α_i^j y su adjunto β_j^i).

Cálculo del adjunto β_j^i de α_i^j. — Calculemos primero β_1^1: se tiene

$$\beta_1^1 = \sum_{p \in S_n, p(1)=1} \varepsilon(p)\alpha_2^{p(2)} \ldots \alpha_i^{p(i)} \ldots \alpha_n^{p(n)}$$

pero hay una biyección entre la parte de $\mathcal{S}_n$ descrita por p tal que $p(1) = 1$ y $\mathcal{S}_{n-1}$: si p' es un elemento de $\mathcal{S}_{n-1}$ operando sobre $\{2, \ldots, n\}$, tal que $p'(i) = p(i)$ $(2 \leq i \leq n)$, los conjuntos ordenados $\{1, p(2), \ldots, p(n)\}$ y $\{p'(2), \ldots, p'(n)\}$ presentan el mismo número de inversiones, de donde con las notaciones precedentes

$$\varepsilon(p) = \varepsilon(p')$$

$$\beta_1^1 = \sum_{p' \in S_{n-1}} \varepsilon(p')\alpha_2^{p'(2)} \ldots \alpha_i^{p'(i)} \ldots \alpha_n^{p'(n)}$$

con $p'(i) \in [2, n]$. β_1^1 es, pues, el determinante de la matriz A_1^1 obtenida suprimiendo en A la primera fila y la primera columna (§ 154, *a*), luego

$$\beta_1^1 = \det (A_1^1).$$

Para calcular β_j^i mediante $i - 1$ transposiciones de columnas consecutivas y $j - 1$ transposiciones de filas consecutivas, llevemos la columna i al primer lugar y la fila j al primer lugar con ello obtenemos una matriz $A' = (\alpha_i'^j)$, donde $\alpha_1'^1 = \alpha_i^j$; consideremos $A_1'^1$ es la matriz A_j^i obtenida suprimiendo en A la i-ésima columna y j-ésima fila, el adjunto $\beta_1'^1$ de $\alpha_1'^1$ es tal que

$$\beta'^1_1 = \det A'^1_1 = \det A_j^i$$

pero como

$$\det A' = (-1)^{i-1+j-1} \det A = (-1)^{i+j} \det A$$

tendremos

$$(2) \qquad \beta_j^i = (-1)^{i+j} \beta'^1_1 = (-1)^{i+j} \det A_j^i$$

(ATENCIÓN: En $(-1)^{i+j}$, $i + j$ es un exponente.)

TEOREMA Y DEFINICIÓN.—*En el desarrollo respecto a los elementos de una columna o de una línea del determinante de la matriz* $A = (\alpha_i^j)$, *el coeficiente* β_j^i *de* α_i^j, *llamado* adjunto *de* α_i^j, *es igual a* $(-1)^{i+j} \det A_j^i$, *siendo* A_j^i *la matriz obtenida suprimiendo en A la columna i y la fila j. El* det A_j^i *se llama el* menor *asociado al elemento* α_i^j.

ATENCIÓN: No confundir *menor* y *adjunto* asociado a α_i^j.

Para los elementos de la diagonal principal $(-1)^{2i} = 1$, luego $\beta_i^i = \det A_i^i$.

b) Aplicación al cálculo de los determinantes

1. La fórmula (1) o (1') conduce el cálculo de un determinante de orden n al cálculo de un determinante de orden $n - 1$, luego en definitiva al cálculo de determinantes de orden 3 o 2; por ejemplo,

$$\begin{vmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{vmatrix} = a_1^1 \begin{vmatrix} a_2^2 & a_3^2 \\ a_2^3 & a_3^3 \end{vmatrix} - a_1^2 \begin{vmatrix} a_2^1 & a_3^1 \\ a_2^3 & a_3^3 \end{vmatrix} + a_1^3 \begin{vmatrix} a_2^1 & a_3^1 \\ a_2^2 & a_3^2 \end{vmatrix}.$$

OBSERVACION

Antes de hacer estos desarrollos respecto a una columna (o una fila) es conveniente hacer aparecer los ceros reemplazando una columna (o una fila) por una combinación lineal de las columnas (o las filas), el multiplicador en la línea reemplazada es igual a 1 (§ 168, *b*, propiedad 4).

2. Siendo A la matriz (α_i^j) cuadrada de orden n, desarrollando el determinante siguiente con relación a la última columna se ve que cualesquiera que sean los escalares λ_i $(1 \leq i \leq n)$

$$\begin{vmatrix} \alpha_1^1 & \cdots & \alpha_n^1 & 0 \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \alpha_1^n & \cdots & \alpha_n^n & 0 \\ \lambda_1 & \cdots & \lambda_n & 1 \end{vmatrix} = \det A$$

la matriz cuyo determinante ha sido escrito por la izquierda ha sido obtenido "orlando" A con una $(n+1)$-ésima columna y una $(n+1)$-ésima fila.

De una manera más general por inducción tendremos

$$\begin{vmatrix} \alpha_1^1 & \cdots & \alpha_n^1 & 0 & \cdots & \cdots & 0 \\ \cdot & & \cdot & \cdot & & & \cdot \\ \cdot & & \cdot & \cdot & & & \cdot \\ \cdot & & \cdot & \cdot & & & \cdot \\ \alpha_1^n & \cdots & \alpha_n^n & 0 & & & \cdot \\ \lambda_1^{n+1} & \cdots & \lambda_n^{n+1} & 1 & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot & \vdots & 1 & 0 \\ \lambda_1^p & \cdots & \lambda_n^p & \cdots & \cdots & \lambda_{p-1}^p & 1 \end{vmatrix} = \det A.$$

Todos los elementos de las columnas $(n+1)$ a p $(p > n)$ situadas encima de 1 de la diagonal principal son nulos, mientras que los elementos de las líneas $(n+1)$ a p situados a la izquierda de los de la diagonal principal son cualesquiera. De donde: *Todo escalar igual a un determinante de orden n puede escribirse en la forma de un determinante de orden $p > n$.*

Esta observación permite escribir el producto de dos determinantes de órdenes respectivos n y n' bajo la forma de un determinante de orden sup (n, n').

3. Sea A una matriz triangular, si A es triangular superior, ${}^tA$ es triangular inferior, como $\det A = \det {}^tA$ podemos suponer A triangular superior, desarrollando A con relación a la primera columna obtenemos

$$\det A = \begin{vmatrix} \alpha_1^1 & \alpha_2^1 & \cdots & \alpha_{n-1}^1 & \alpha_n^1 \\ 0 & \alpha_2^1 & \cdots & \alpha_{n-1}^2 & \alpha_n^2 \\ 0 & 0 & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \alpha_{n-1}^{n-1} & \alpha_n^{n-1} \\ 0 & 0 & \cdots & 0 & \alpha_n^n \end{vmatrix} = \alpha_1^1 \begin{vmatrix} \alpha_2^2 & \alpha_2^3 & \cdots & \alpha_n^2 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \cdots & \alpha_n^n \\ 0 & \alpha_3^3 & \cdots & \alpha_n^n \end{vmatrix}$$

y por inducción

$$\det A = \alpha_1^1 \alpha_2^2 \cdots \alpha_n^n,$$

TEOREMA. — *El determinante de una matriz triangular A es igual al producto de los elementos diagonales de A.*

171. Determinantes cuyos elementos pertenecen a un anillo

Sea E un conjunto provisto de una adición y de una multiplicación y $M = (\alpha_i^j)$ una matriz cuadrada de orden n con elementos en E (ver § 159), pondremos por definición

$$(1) \qquad \det M = \det (\alpha_i^j) = \sum_{p \in S_n} \varepsilon(p) \alpha_1^{p(1)} \dots \alpha_n^{p(n)}.$$

Esta definición formal, como la de las operaciones de las matrices con elementos en E (§ 159), no tendrá interés más que si ciertas propiedades y métodos de cálculo relativos a los determinantes de elementos en un cuerpo conmutativo se aplican a los determinantes que acabamos de definir.

Este será siempre el caso si E = A es un anillo íntegro; en efecto, este último puede siempre estar considerado como un subanillo de un cuerpo de fracciones: todas las propiedades relativas a las sumas, sustracciones, multiplicaciones permanecen válidas: las enunciadas en los §§ 168, 169 y 170, excepto naturalmente las relativas a la inversión de M (ver § 173, observación 2).

Análogamente si E = A es un anillo conmutativo unitario, si los cálculos no hacen intervenir la división en A (salvo para los elementos inversibles de A): es suficiente sustituir la noción de espacio vectorial sobre K por la de A-módulo libre de tipo finito (ver § 159 y capítulo 7, ej. 179).

III. Primeras aplicaciones de los determinantes

Una de las aplicaciones más importantes de los determinantes es el cálculo efectivo de las soluciones de un sistema de ecuaciones lineales, la daremos en el capítulo siguiente. Estudiamos aquí tres aplicaciones relativas a los espacios vectoriales y a las matrices.

172. Orientación de un espacio vectorial real de dimensión finita

En todo este párrafo, entendemos por base de E, de dimensión n sobre **R**, un n-étuple de E^n formado de vectores linealmente independientes y no una simple parte $\{a_1, \dots, a_n\}$ formada por n vectores linealmente independientes de E.

Consideremos dos bases (a_i) y (b_j) de un espacio vectorial de E de dimensión n sobre **R**, consideremos la relación entre (a_i) y (b_j) definida por

$$\det_{(a_i)} (b_1, \dots, b_n > 0$$

vamos a demostrar que es una *relación de equivalencia* definida sobre el conjunto de las bases de E.

Es *reflexiva*, pues por definición

$$\det_{(a_i)}(a_1, \ldots, a_n) = 1 > 0.$$

Es *simétrico*, pues siendo P la matriz de paso de la base (a_i) a la base (b_j) tenemos (§ 160)

$$P = M(\mathrm{id}_E, (b_j), (a_i)), \qquad P^{-1} = M(\mathrm{id}_E, (a_i), (b_j))$$

luego

$$\det P = \det_{(a_i)}(b_1, \ldots, b_n), \qquad \det P^{-1} = \det_{(b_j)}(a_1, \ldots, a_n).$$

La relación $\det P \det P^{-1} = 1$ muestra bien que estos dos determinantes son estrictamente positivos simultáneamente.

Finalmente es *transitiva*: sean tres bases (a_i), (b_j), (c_k) de E tendremos (§ 157, *a*)

$$M(\mathrm{id}_E, (c_k), (a_i)) = M(\mathrm{id}_E, (c_k), (b_j)\; M(\mathrm{id}_E, (b_j), (a_i))$$

de donde

$$\det_{(a_i)}(c_1, \ldots, c_n) = \det_{(b_j)}(c_1, \ldots, c_n) \det_{(a_i)}(b_1, \ldots, b_n)$$

en consecuencia, si los dos determinantes del segundo miembro son estrictamente positivos, lo es también el determinante del primer miembro:

Teorema y definición. — *Dadas dos bases (a_i) y (b_j) de un espacio vectorial* E *de dimensión n sobre* **R** *la relación entre estas bases definida por*

$$\det_{(a_i)}(b_1, \ldots, b_n) > 0$$

es una relación de equivalencia. Se dice que las bases (a_i) y (b_j) que verifican esta relación tienen la misma orientación.

Escogida una base B_0 de E, vemos que esta relación define una partición del conjunto de las bases de E en dos clases:

— Las bases B que tienen la misma orientación que B_0.

— Las bases B′ que tienen la *orientación opuesta* a la de B_0 (es decir, tal que el determinante de las coordenadas de los vectores de B′ con relación a B_0 es estrictamente negativo).

Escoger una de estas clases es por definición orientar el espacio E: todas las bases que pertenecen a esta clase se llaman *directas* o *positivamente orientadas,* las bases pertenecientes a la otra clase se llaman *inversas* o *negativamente orientadas.*

Para orientar el espacio en general se escoge una base (a_i) y se califica de directas las bases que tienen la misma orientación.

OBSERVACIONES

1. Repitamos que por base B entendemos un n-étuple $(a_1, \ldots, a_n)$ de E^n y no una parte $\{a_1, \ldots, a_n\}$ de E, así las dos bases $(a_1, \ldots, a_n)$ y $(a_{p(1)}, \ldots, a_{p(n)})$ son de la misma orientación si y sólo si la permutación p es par.

2. En $\mathbf{R}^n$ existe una base canónica $(e_1, \ldots, e_n)$ es natural admitir que es directa. Pero en un espacio vectorial E de dimensión n sobre $\mathbf{R}$ no hay en general ninguna base en que la elección se imponga de manera natural respecto a las demás. En consecuencia, la elección de una base calificada de directa es arbitraria.

Por otra parte, la elección de bases directas en el espacio «concreto» (de dos o tres dimensiones) mediante la ayuda de nociones de derecha y de izquierda procede de la Física y no de las Matemáticas. Por ejemplo, veremos que los polinomios de grado inferior o igual a 1 de coeficientes reales describen un espacio vectorial de dimensión 2 sobre $\mathbf{R}$ (§ *187*) del que una base es, por ejemplo, $a_1 = x - 1$, $a_2 = x + 1$; decir que a_2 está a la izquierda de a_1 no tiene ningún sentido.

173. Cálculo de la inversa A^{-1} de una matriz cuadrada A inversible

Sea $A = (\alpha_i^j)$ un elemento de $M_n(K)$; designemos por $x_1, x_2, \ldots, x_n$ sus vectores columnas y por β_j^i el adjunto de α_i^j en el desarrollo de det A.

Consideremos la suma

$$\gamma_{i'}^i = \sum_{j=1}^n \beta_j^i \alpha_{i'}^j = \beta_1^i \alpha_{i'}^1 + \ldots + \beta_j^i \alpha_{i'}^j + \ldots + \beta_n^i \alpha_{i'}^n.$$

Si $i = i'$ según la igualdad (1) del § 170, $\gamma_i^i = \det A$.

Si $i \neq i'$ siempre según la misma fórmula $\gamma_{i'}^i$ es el desarrollo de un determinante de orden n con relación a la columna de número i, siendo la columna de índice $k \neq i$ igual a x_k y la columna de índice i igual a x'_i: este determinante tiene, pues, dos columnas idénticas, la de número i' y la de número i, y es, en consecuencia, nulo; tendremos entonces ($\delta_{i'}^i$, símbolo de KRONECKER)

$$\text{(1)} \qquad \sum_{j=1}^n \beta_j^i \alpha_{i'}^j = \delta_{i'}^i \det A$$

tendríamos razonando igualmente con las líneas de det A

$$\text{(2)} \qquad \sum_{i=1}^n \beta_j^i \alpha_i^{j'} = \delta_j^{j'} \det A.$$

Poniendo

$$B = (\beta_j^i), \qquad L = BA = (\lambda_i^j), \qquad M = AB = (\mu_i^j)$$

obtendremos

$$\text{(1')} \qquad \lambda_i^j = \sum_{k=1}^n \beta_k^j \alpha_i^k = \delta_i^j \det A$$

reemplazando en (1) i', i, j, respectivamente, por i, j, k; igualmente la fórmula (2) nos da

$$\text{(2')} \qquad \mu_i^j = \sum_{k=1}^n \alpha_k^j \beta_i^k = \delta_i^j \det A$$

luego

(3) $$BA = AB = I_n \text{ (det A)}$$

siendo K un cuerpo, todo elemento no nulo de K es inversible; volvemos a encontrar que A es inversible si y sólo si det $A \neq 0$ y obtenemos

(4) $$A^{-1} = [(\det A)^{-1}]B.$$

En $B = (\beta_j^i)$, j es el índice de la *columna* e i el de la *fila*; si designamos por *adjunto* de A la matriz $B' = {}^tB$, $\beta_i'^j = \beta_j^i$ adjunto de α_i^j tiene en ${}^tB$ el mismo lugar que α_i^j en A: aquí está el interés de la introducción de la matriz adjunta ${}^tB$ de A, de donde:

TEOREMA. — *La inversa* A^{-1} *de una matriz inversible* A *de* $M_n(K)$ *se obtiene multiplicando por* $(\det A)^{-1}$ *la transpuesta de la matriz adjunta de* A.

OBSERVACIONES

1. Veremos en el capítulo 10 (observación final del § 176) que la resolución de un sistema lineal asociado a A inversible (sistema de CRAMER), proporciona algunas veces un cálculo de A^{-1} más simple que el precedente.

2. Si K no es un cuerpo conmutativo, sino un anillo conmutativo unitario, la fórmula (3) es también válida (ver § 171) y tenemos, pues, el resultado siguiente: A, *matriz cuadrada de elementos en un anillo conmutativo unitario es inversible si y sólo si* det A *es inversible en el anillo.*

EJERCICIO

Siendo A un elemento de $M_n(K)$, demostrar que

(3′) $$(\det A)(\det B) = (\det A)^n$$

deducir que si $\det A \neq 0$, $\det B = (\det A)^{n-1}$.

(Observar que la fórmula (3′) comparada con la fórmula (3) pone en evidencia el carácter n-lineal del determinante.)

174. Determinación del rango de una matriz

Sea $A = (\alpha_i^j)$ $(1 \leq i \leq m,\ 1 \leq j \leq n)$ una matriz de tipo (m, n) sobre **K**, su rango r es la dimensión del subespacio de K^n engendrado por sus vectores columnas $x_1, x_2, \ldots, x_m$ con (siendo $e_1, \ldots, e_n$ la base canónica de K^n)

$$i = 1, 2, \ldots, m \qquad x_i = \sum_{j=1}^{n} \alpha_i^j e_j.$$

Hemos visto que si $m = n$, rg (A) $= n$ si y sólo si det (A) $\neq 0$.

Vamos a generalizar este resultado asociando el rango de A al orden máximo de una matriz cuadrada regular extraída de A.

Llamaremos *menor* de la matriz A todo determinante de una matriz cuadrada extraída de A.

Demostremos primero una propiedad preliminar.

a) **Condición para que un vector pertenezca a un subespacio de** E

Sea $x_1, x_2, \ldots, x_s$, s vectores de un espacio vectorial E de dimensión n sobre K, expresado en una base $a_1, \ldots, a_n$

$$x_i = \sum_{j=1}^{n} \alpha_i^j a_j \qquad (i = 1, 2, \ldots, s).$$

Supongamos que la matriz de las coordenadas de estos vectores (que es de tipo (s, n)) posee un menor de orden s no nulo. Cambiando si fuera preciso la numeración de los vectores $a_1, \ldots, a_n$ podemos suponer que

$$\Delta = \begin{vmatrix} \alpha_1^1 & \cdots & \alpha_s^1 \\ \vdots & & \vdots \\ \alpha_1^s & \cdots & \alpha_s^s \end{vmatrix} \neq 0.$$

Si $s = n$ los vectores son independientes. Estudiemos el caso $s < n$. A todo vector $x = \alpha^1 a_1 + \ldots + \alpha^n a_n$ de E, hagamos corresponder el vector de E definido por

$$f(x) = \alpha^1 a_1 + \ldots + \alpha^s a_s.$$

La aplicación f es la proyección de E sobre el subespacio E′ engendrado por $a_1, \ldots, a_s$ paralelamente al subespacio engendrado por $a_{s+1}, \ldots, a_n$; f es, pues, un endomorfismo de E (ver § 139, ej. 1), $\Delta \neq 0$ implica que los vectores $f(x_1), \ldots, f(x_s)$ son independientes; estos vectores, perteneciendo a E′ de dimensión s, determinan una base de E′. Ahora bien, $f(x)$ pertenece a E′; existen, pues, escalares $\lambda^1, \ldots, \lambda^s$ tales que

$$(1) \qquad f(x) = \lambda^1 f(x_1) + \ldots + \lambda^s f(x_s).$$

Pero $f(x_1), \ldots, f(x_s)$ siendo independientes en E′ lo son en E y $x_1, \ldots, x_s$ lo son igualmente en E (ver § 133, *b*); $x_1, \ldots, x_s$ independientes engendran un subespacio F, de dimensión s, de E.

Busquemos en qué condiciones un vector x cualquiera de E pertenece a F, para esto deberán existir escalares $\mu^1, \ldots, \mu^s$ tales que

$$(2) \qquad x = \mu^1 x_1 + \ldots + \mu^s x_s.$$

En este caso tendremos

$$(2') \qquad f(x) = \mu^1 f(x_1) + \ldots + \mu^s f(x_s),$$

pero según (1), al ser única la descomposición de un vector sobre una base se deberá tener que $\lambda^i = \mu^i$ para todo i de $[1, s]$. Por otra parte, la igualdad (2) es equivalente al sistema (3) (después de sustituir μ^i por λ^i).

$$(3) \qquad \alpha^k = \lambda^1 \alpha_1^k + \ldots + \lambda^s \alpha_s^k \qquad (k = 1, 2, \ldots, n).$$

Según (1) estas igualdades se verifican para $1 \leq k \leq s$, *finalmente x pertenecerá al subespacio* F *engendrado por $x_1, \ldots, x_s$ si y sólo si las igualdades* (3) *se verifican para $s + 1 \leq k \leq n$.*

Para cada valor k de $[s+1, n]$, hagamos corresponder a x el vector

(4) $$f_k(x) = \alpha^1 a_1 + \dots + \alpha^s a_s + \alpha^k a_k = f(x) + \alpha^k a_k$$

y consideremos el determinante Δ_k de los vectores $f_k(x_1), \dots, f_k(x_s), f_k(x)$ con relación a la base $a_1, a_2, \dots, a_s, a_k$ del subespacio al que pertenecen

$$\begin{aligned}\Delta_k &= \det\,[f_k(x_1), \dots, f_k(x_s), f_k(x)] \\ &= \det\,[f_k(x_1), \dots, f_k(x_s), f_k(x) - \lambda^1 f_k(x_1) \dots - \lambda^s f_k(x_s)]\end{aligned}$$

según (4) y (1)

$$\Delta_k = \begin{vmatrix} \alpha_1^1 & \dots & \alpha_s^1 & \alpha^1 \\ \vdots & & \vdots & \vdots \\ \alpha_1^s & \dots & \alpha_s^s & \alpha^s \\ \alpha_1^k & \dots & \alpha_s^k & \alpha^k \end{vmatrix} = \begin{vmatrix} \alpha_1^1 & \dots & \alpha_s^1 & 0 \\ \vdots & & \vdots & \vdots \\ \alpha_1^s & \dots & \alpha_s^s & 0 \\ \alpha_1^k & \dots & \alpha_s^k & \alpha^k - \lambda^1\alpha_1^k \dots - \lambda^s\alpha_s^k \end{vmatrix} \quad (k = s+1, \dots, n)$$

desarrollando con relación a la última columna el determinante escrito en la segunda forma, obtenemos

$$\Delta_k = \Delta(\alpha^k - \lambda^1\alpha_1^k \dots - \lambda^s\alpha_s^k) \qquad (k = s+1, \dots, n).$$

Luego no siendo Δ nulo, las igualdades (3) se verifican si y sólo si $\Delta_k = 0$ para todo k de $[s+1, n]$.

Los determinantes Δ_k (escritos en la primera forma) se llaman los *determinantes característicos* de x (con relación a los vectores independientes $f(x_1), \dots, f(x_s)$). Se obtienen orlando el determinante Δ de las s primeras coordenadas de $x_1, \dots, x_s$ por la derecha con las coordenadas del mismo número de x y por abajo por las k-ésimas coordenadas de $x_1, \dots, x_s, x$; de donde:

LEMA. — *Si el determinante Δ de las s primeras coordenadas de s vectores $x_1, \dots, x_s$ de un espacio de dimensión n es no nulo, estos vectores son independientes. Un vector x pertenece al subespacio engendrado por $x_1, \dots, x_s$, si y sólo si sus $n-s$ determinantes característicos son nulos.*

b) Determinación del rango de una matriz

Sea una matriz $A = (\alpha_i^j)$ de tipo (m, n) sobre K, designemos sus vectores columnas por $x_1, x_2, \dots, x_m$ pertenecen a un espacio E de dimensión n sobre K.

Sea s el mayor entero tal que existe un menor de A de orden s no nulo,

$$s \leq \inf(m, n).$$

Sea Δ uno de estos menores, según la primera parte del lema precedente los vectores cuyas coordenadas pertenecen a Δ son independientes y engendran un subespacio F de E de dimensión s; sean (x_i) (i perteneciendo a S parte de s elementos de $M = [1, m]$) estos vectores. Cada uno de los $m - s$ vectores x_j (j perteneciendo al complementario de S con relación a M) pertenece a F: en efecto, todas las características de x_j son nulas siendo menores de orden $s+1$ extraídos de la matriz: entonces el subespacio engen-

drado por $x_1, x_2, \ldots, x_m$ es idéntico a F de dimensión s: s es, pues, el rango de los vectores (x_i) $(1 \leq i \leq m)$, es decir, el rango de r de la matriz A es s (ver §§ 138 y 162, *a*), de donde:

Teorema. — *El rango de una matriz* A *es igual al orden máximo de un menor no nulo extraído de la matriz* A.

OBSERVACION

El orden máximo de un menor no nulo es evidentemente el mismo para A y ${}^tA$; hemos demostrado, pues, de un nuevo modo que rg (A) = rg (${}^tA$).

El corolario siguiente resume las propiedades obtenidas en este párrafo, y en los §§ 143 y 162:

Corolario. — *Dada una aplicación lineal f de un espacio vectorial* E *en un espacio vectorial* F, *los dos de dimensiones finitas, todos los apartados siguientes son iguales* (A = M(f)):

a) dim f(E) = rg (f) = rg (${}^tf$).
b) rg (A) = rg (${}^tA$).
c) *Número máximo de vectores columnas (o de vectores filas) de* A, *independientes.*
d) *Orden máximo de un menor no nulo extraído de* A.

EJERCICIOS

1. Hallar el rango de las matrices

$$\begin{pmatrix} 2 & -1 & 4 & -4 \\ 3 & 2 & -3 & 17 \\ 5 & -3 & 8 & -10 \end{pmatrix} \quad \begin{pmatrix} 3 & -1 & 6 \\ -1 & 1 & -1 \\ 7 & -3 & 11 \\ -4 & 2 & -5 \end{pmatrix}.$$

2. Siendo p, q, r reales, determinar el rango de la matriz

$$\begin{pmatrix} 0 & -r & q \\ r & 0 & -p \\ -q & p & 0 \end{pmatrix}.$$

Ejercicios

Salvo indicación contraria, los elementos de los determinantes considerados se toman de un cuerpo conmutativo cualquiera.

232. **Demostrar**

$$\begin{vmatrix} 0 & a & b \\ a & 0 & c \\ b & c & 0 \end{vmatrix} = 2abc.$$

233. **Demostrar**

$$\begin{vmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2a \\ 2c & 2c & c-a-b \end{vmatrix} = (a+b+c)^3.$$

234. Demostrar

$$\begin{vmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & c^2 & b^2 \\ 1 & c^2 & 0 & a^2 \\ 1 & b^2 & a^2 & 0 \end{vmatrix} = \begin{vmatrix} 0 & a & b & c \\ a & 0 & c & b \\ b & b & 0 & a \\ c & c & a & 0 \end{vmatrix} = -(a+b+c)(b+c-a)(c+a-b)(a+b-c).$$

235. K = **R** (o K es un cuerpo de característica $\neq 2$). Demostrar que los dos determinantes

$$\begin{vmatrix} 0 & a \\ -a & 0 \end{vmatrix} \qquad \begin{vmatrix} 0 & a & b & c \\ -a & 0 & r & -q \\ -b & -r & 0 & p \\ -c & q & -p & 0 \end{vmatrix}$$

son los cuadros de los elementos de K. (Generalización: V. ej. 261).

236. Se pone (K = **R** o **C**, $j^3 = 1$)

$$P(x,\ y,\ z) = \begin{vmatrix} x & y & z \\ y & z & x \\ z & x & y \end{vmatrix}.$$

a) Demostrar que

$$\begin{aligned} P(x,y,z) = 3xyz - (x^3+y^3+z^3) &= -(x+y+z)(x^2+y^2+z^2-yz-zx-xy) \\ &= -(x+y+z)(x+jy+jz)(x+j^2y+j^2z). \end{aligned}$$

(Generalización: V. Ej. 259).

b) Siendo (x, y, z) y (x', y', z') dos elementos dados de K^3 demostrar que existe (x'', y'', z'') de K^3 tal que

$$P(x,y,z)\ \ P(x',y',z') = P(x'',y'',z'').$$

237. K = **R** se pone

$$P(x,y,z,t) = \begin{vmatrix} x & y & z & t \\ -y & x & -t & z \\ -z & t & x & -y \\ -t & -z & y & x \end{vmatrix}.$$

a) Demostrar que $P(x, y, z, t) = (x^2 + y^2 + z^2 + t^2)^2$.

b) Siendo (x, y, z, t) y (x', y', z', t') dos elementos dados de $\mathbf{R}^4$ demostrar que existe (x'', y'', z'', t'') de $\mathbf{R}^4$ tal que

$$P(x, y, z, t)P(x', y', z', t') = P(x'', y'', z'', t'').$$

Deducir una propiedad de los enteros productos de dos sumas de 4 cuadros.

238. $\mathbf{K} = \mathbf{R}$. Demostrar la relación

$$(x^2 + y^2)(x'^2 + y'^2) = (xx' + yy')^2 + (xy' - yx')^2$$

calculando

$$\begin{vmatrix} x & -y \\ y & x \end{vmatrix} \qquad \begin{vmatrix} x' & y' \\ -y' & x' \end{vmatrix};$$

239. a, b, c son reales, se pone $a + b + c = 2p$ demostrar

$$\begin{vmatrix} 1 & \cos c & \cos a \\ \cos c & 1 & \cos b \\ \cos b & \cos a & 1 \end{vmatrix} = 4 \operatorname{sen} p \operatorname{sen}(p - a) \operatorname{sen}(p - b) \operatorname{sen}(p - c).$$

240. a, b, c son reales, demostrar

$$\begin{vmatrix} 1 & \operatorname{sen} a & \cos a \\ 1 & \operatorname{sen} b & \cos b \\ 1 & \operatorname{sen} c & \cos c \end{vmatrix} = \operatorname{sen}(b - c) + \operatorname{sen}(c - a) + \operatorname{sen}(a - b)$$
$$= -4 \operatorname{sen} \frac{b-c}{2} \operatorname{sen} \frac{c-a}{2} \operatorname{sen} \frac{a-b}{2}.$$

241. a, b, c son reales, demostrar

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & \cos c & \cos b \\ 1 & \cos c & 1 & \cos a \\ 1 & \cos b & \cos a & 1 \end{vmatrix} = -4 \operatorname{sen}^2 \frac{a}{2} \operatorname{sen}^2 \frac{b}{2} \operatorname{sen}^2 \frac{c}{2}.$$

242. Calcular

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{vmatrix}.$$

Más generalmente, demostrar que si en un determinante de orden n se tiene

$$(j = 1, 2, \ldots, n) \qquad a^j_{i'} = a^j_i + r, \qquad a^j_{i''} = a^j_{i'} + r$$

el determinante es nulo (r dado, i, i', i'' son tres índices de columnas distintas 2 a 2).

En los ejercicios siguientes (243 *al* 247) *se considerará* det A, ($A = (\alpha^j_i)$) *como un polinomio de* n^2 *indeterminadas* α^j_i *y se utilizará las propiedades siguientes* (V. § 195):

1) $f(X, Y, Z, \ldots)$ es divisible por $X - Y$ si, y sólo si $f(X, X, Z, \ldots) = 0$.

2) Si $f(X_1, \ldots, X_m)$ es divisible por $X_i - X_j (1 \leq i < j \leq n)$ entonces f es divisible por el producto de estos polinomios (V. § 195, corolario del teorema 18 y ejercicio 5).

243. Determinante de *Vandermonde*. Se pone

$$D(a_0, \ldots, a_n) = \begin{vmatrix} 1 & a_0 & (a_0)^2 & \ldots & (a_0)^n \\ 1 & a_1 & (a_1)^2 & \ldots & (a_1)^n \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & a_n & (a_n)^2 & \ldots & (a_n)^n \end{vmatrix}.$$

(el índice superior es un exponente, D es de orden $n + 1$).

a) Demostrar que D es homogéneo de grado total $\dfrac{n(n+1)}{2}$.

b) Demostrar que D es divisible por $a_i - a_j$ $(i \neq j)$.

c) Deducir de las propiedades precedentes que

$$D(a_0, \ldots, a_n) = \prod_{0 \leq i < j \leq n} (a_j - a_i).$$

244. Calcular

$$\begin{vmatrix} 1 & a & a^2 & a^4 \\ 1 & b & a^3 & b^4 \\ 1 & b^2 & b^3 & c^4 \\ 1 & c^2 & c^3 & d^4 \end{vmatrix}.$$

245. Calcular

$$\begin{vmatrix} b+c & bc-1 & (b^2+1)(c^2+1) \\ c+a & ca-1 & (c^2+1)(a^2+1) \\ a+b & ab-1 & (a^2+1)(b^2+1) \end{vmatrix}.$$

246. Calcular

$$\begin{vmatrix} b+c & c+a & a+b \\ b^2+c^2 & c^2+a^2 & a^2+b^2 \\ b^3+c^3 & c^3+a^3 & a^3+b^3 \end{vmatrix}.$$

247. Siendo $a_0, \ldots, a_n$ números reales calcular el determinante cuya fila k $(0 \leq k \leq n)$ es

$$1 \ \cos a_k \ \cos 2a_k \ldots \cos na_k.$$

(Se expresará $\cos pa_k = P(\cos a_k)$ y se observará que para calcular el determinante sólo es necesario conocer el monomio de más alto grado, poniendo $\cos a_k = x_k$ nos encontramos con un determinante de VANDERMONDE.)

En los ejercicios siguientes (248 *al* 252) *se hallará una relación de recurrencia entre* D_n y D_{n-1} *o entre* D_n, D_{n-1}, D_{n-2}.

248. Calcular

$$D_n = \begin{vmatrix} a_0 & a_1 & a_2 & \ldots & a_{n-1} & a_n \\ -1 & x & 0 & \ldots & 0 & 0 \\ 0 & -1 & x & \ldots & 0 & 0 \\ \cdot & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ 0 & 0 & 0 & \ldots & x & 0 \\ 0 & 0 & 0 & \ldots & -1 & x \end{vmatrix}$$

$(D_n = xD_{n-1} + a_n)$.

249. Calcular D_n, de orden n, cuyos elementos son todos iguales a 1 salvo los de la diagonal principal que son iguales a $1 + x$.

250. Calcular D_n, de orden n, cuyos elementos son todos iguales a x excepto los de la diagonal principal que son iguales a -1.

251. Calcular D_n, de orden n

$$D_n = \begin{vmatrix} 1+x^2 & -x & 0 & \dots & 0 & 0 & 0 \\ -x & 1+x^2 & -x & \dots & 0 & 0 & 0 \\ 0 & -x & 1+x^2 & \dots & 0 & 0 & 0 \\ \vdots & & & & & & \\ 0 & 0 & 0 & \dots & -x & 1+x^2 & -x \\ 0 & 0 & 0 & \dots & 0 & -x & 1+x^2 \end{vmatrix}$$

(Se hallará una relación entre D_n, D_{n-1}, D_{n-2}, a continuación habiendo puesto $\Delta_n = D_n - D_{n-1}$ se hallará una relación entre Δ_n y Δ_{n-1}.)

252. Calcular $D_n (n \geq m - 1$, C_n^p coeficiente binomial)

$$D_n = \begin{vmatrix} 1 & C_n^1 & C_n^2 & \dots & C_n^{m-1} \\ 1 & C_{n+1}^1 & C_{n+1}^2 & \dots & C_{n+1}^{m-1} \\ \vdots & & & & \\ 1 & C_{n+m-1}^1 & C_{n+m-1}^2 & \dots & C_{n+m-1}^{m-1} \end{vmatrix}$$

$(D_m = D_{m-1})$.

253. Se pone (n y k enteros naturales)

$$D(n, k) = \begin{vmatrix} 1^k & 2^k & \dots & n^k \\ 2^k & 3^k & \dots & (n+1)^k \\ \vdots & & & \\ n^k & (n+1)^k & \dots & (2n-1)^k \end{vmatrix}$$

a) Calcular $D(n, 1)$ para $n = 1, 2, 3$ (V. ej. 242).

b) Demostrar que $D(n, 2)$ es nulo para $n > 3$.

c) Demostrar que $D(n, k)$ es nulo para $n > k + 1$.

254. Si (a_k) y (b_k) son dos familias de elementos de $K (1 \leq k \leq n)$ calcular el determinante D_n cuya k-ésima fila es

$$b_k \quad b_k \dots b_k \quad a_k + b_k \quad b_k \dots b_k$$

donde $a_k + b_k$ está sobre la diagonal principal.
(Calcular una relación entre D_n y D_{n-1}.)

255. Si (a_k) y (b_k) son dos familias de elementos de $K (1 \leq k \leq n)$ tales que para todo (i, j) de $[1, n]^2$, $a_i + b_j \neq 0$, calcular el determinante

$$D_n = \det\left(\frac{1}{a_i + b_j}\right).$$

(Se pondrá $P = \prod_{i=1}^{n} \prod_{j=1}^{n} (a_i + b_j)$, demuéstrese que $D_n\Delta$ es nulo si $i' \neq i$, $a_i = a_{i'}$ y $j' \neq j$, $b_j = b_{j'}$ y se aplicará la observación precedente al ejercicio 243.)

256*. Siendo (r_k) una familia de n elementos de K se pone

$$f(x) = (r_1 - x)(r_2 - x) \dots (r_n - x)$$

y se llama D el determinante cuya k-ésima fila es

$$bb \dots br_k a \dots a$$

con r_k sobre la diagonal principal y $a \neq b$. Demostrar que

$$D = [af(b) - bf(a)](b - a)^{-1}$$

(restando la última fila de todas las demás se demostrará que

$$D = f(a) + aD_1$$

se demostrará en seguida que $D = f(b) + bD_2$ y que $D_1 = D_2$).

257. Sea $A(t)$ una matriz de orden n cuyos elementos son $\alpha_i^j(t)$, las n^2 funciones α_i^j son funciones reales derivables de la variable real t; de designa por $C_i(t)$ los n vectores columnas de $A(t)$ y se escribe

$$f(t) = \det A(t) = \det (C_1(t), \dots, C_n(t)).$$

a) Demostrar

$$f'(t) = \sum_{k=1}^{n} \det (C_1(t), \dots, C_{k-1}(t), C'_k(t), C_{k+1}(t), \dots, C_n(t))$$

b) En el ejercicio 249 se pone $f(x) = D_n$ (x real) calcular $f'(x)$, deducir D_n de ella.

258*. Se tiene $\omega = \cos 2\pi/n + i \operatorname{sen} 2\pi/n$, $\alpha = \cos \pi/n + i \operatorname{sen} \pi/n$ y se considera la la matriz cuadrada X de orden n cuyo elemento general es

$$x_{pq} = \omega^{(p-1)(q-1)} \qquad (1 \leq p \leq n,\ 1 \leq q \leq n)$$

y se pone $D = \det X$.

a) Calcular X^2 y mostrar que $D^2 = (-1)^{\frac{(n-1)(n-2)}{2}} n^n$.

¿En qué caso D es real?

b) Probar que

$$D = \prod (\alpha^{2l} - \alpha^{2k}) = \left[\prod (\alpha^{k+l})\right] \left[\prod 2i \operatorname{sen} \frac{l-k}{n} \pi\right]$$

donde los productos se obtienen al variar k y l en $1 \leq k < l \leq n-1$;

c) Demostrar que

$$|D| = \prod 2 \operatorname{sen} \frac{l-k}{n} \pi = n^{n/2},$$

se pone

$$P = \prod \alpha^{k+l}$$

donde los productos están extendidos a $1 \leq k < l \leq n-1$ pruébese que

$$P = i^{\frac{(n-1)(3n-2)}{2}}$$

deducir el valor de D.

259*. *Determinantes circulantes.*

Si $a_0, a_1, \ldots, a_{n-1}$ son n números complejos se considera las matrices

$$M = \begin{pmatrix} a_0 & a_1 \ldots a_{n-1} \\ a_1 & a_2 \ldots a_0 \\ \vdots & \\ a_{n-1} & a_0 \ldots a_{n-2} \end{pmatrix}; \quad Y = MX, \; Z = XMX$$

donde X es la matriz definida en el ejercicio precedente.
Se pone (p índice de fila, q de columna)

$$M = (m_{pq}), \qquad X = (x_{pq}), \qquad Y = (y_{pq}), \qquad Z = (z_{pq}).$$

a) Demostrar que

$$y_{pq} = \omega^{-(p-1)(q-1)} \sum_{k=0}^{n-1} \omega^{k(q-1)} a_k$$

$$p \neq q \; : \; z_{pq} = 0, \qquad z_{pp} = n \sum_{l=0}^{n-1} \omega^{l(p-1)} a_l$$

(Si h es un entero se preferirá representar por (h) el entero natural verificado

$$(h) \equiv h \pmod{n} \qquad o \leq (h) \leq n-1$$

así $m_{pq} = a_{(p+q-2)}$.)

b) Calcular det Z, de ello deducir

$$\det M = (-1)^{\frac{(n-1)(n-2)}{2}} \prod_{h=0}^{n-1} \left[\sum_{k=0}^{n-1} \omega^{hk} a_k \right]$$

(utilizar el valor de det X^2, ej. 258).

c) Desarrollar la fórmula obtenida para $n = 2, 3, 4$.

260. Si $A = (\alpha_i^j)$ es una matriz de orden n, se designa por β_j^i el adjunto de α_i^j. Se propone generalizar la fórmula obtenida en el ejercicio del § 173. Se escribe $D = \det A$, k es un entero $0 < k < n$.

a) Demostrar la relación siguiente

$$\begin{pmatrix} I_k & O \\ \beta_1^{k+1} \ldots \beta_k^{k+1} & \beta_{k+1}^{k+1} \ldots \beta_n^{k+1} \\ \vdots & \vdots \\ \beta_1^n \ldots \beta_k^n & \beta_{k+1}^n \ldots \beta_n^n \end{pmatrix} A = \begin{pmatrix} \alpha_1^1 \ldots \alpha_k^1 & \alpha_{k+1}^1 \ldots \alpha_n^1 \\ \vdots & \\ \alpha_1^k \ldots \alpha_k^k & \alpha_{k+1}^k \ldots \alpha_n^k \\ O & DI_{n-k} \end{pmatrix}$$

(Utilizar las relaciones (1) y (2) del § 173);

b) Deducir de la relación precedente que, si $D \neq 0$, se tiene

$$\begin{vmatrix} \beta_{k+1}^{k+1} & \cdots & \beta_n^{k+1} \\ \vdots & & \vdots \\ \beta_{k+1}^n & \cdots & \beta_n^n \end{vmatrix} = D^{n-k-1} \begin{vmatrix} \alpha_1^1 & \cdots & \alpha_k^1 \\ \vdots & & \vdots \\ \alpha_1^k & \cdots & \alpha_k^k \end{vmatrix}$$

261. Se supone $\mathbf{K} = \mathbf{R}$ (o $\mathbf{K}$ cuerpo conmutativo de característica $\neq 2$) y se considera el determinante D_{2m} de una matriz antisimétrica de orden par. Demostrar que D_{2m} es el cuadrado de un elemento de $\mathbf{K}$.
(Razonar por recurrencia, siendo el resultado verdadero para $2m = 2$; utilizar el ejercicio precedente con $n = 2m$, $k = 2m - 2$, y el ejercicio del § 168.)

262. Se considera la matriz B, matriz de cuatro bloques, A y A′ son dos matrices cuadradas de orden respectivo n y n'

$$\mathbf{B} = \begin{pmatrix} A & C \\ O & A' \end{pmatrix}$$

demostrar que

$$\det B = (\det A)\,(\det A').$$

(Se pondrá $A = (\alpha_i^j)$, $A' = (\alpha'^j_i)$, $B = (\beta_i^j)$ y se demostrará que entre los términos de det B

$$\varepsilon(p)\beta_1^{p(1)} \ldots \beta_{n+n'}^{p(n+n')}$$

sólo se consideran los términos correspondientes a la permutación p que verifica o bien $p(i) \in \{1, 2, \ldots, n\}$, para $i = 1, 2, \ldots, n$, o bien $p(n + i') \in \{n+1, n+2, \ldots, n+n'\}$, para $i' = 1, 2, \ldots, n'$, mientras que todos lo sdemás son nulos.)

263. Tenemos la matriz cuadrada $A = (A_i^j)$ de orden n en que para $i < j$ la matriz A_i^j es nula y para $i = 1, 2, \ldots, n$ A_i^i es una matriz cuadrada.
Demostrar por recurrencia utilizando el ejercicio precedente que

$$\det A = (\det A_1^1)(\det A_2^2) \ldots (\det A_n^n).$$

264*. Se designa respectivamente por I y por J las partes $\{i_1, \ldots, i_m\}$, $\{j_1, \ldots, j_m\}$ de m elementos de $[1, n]$ $(m \leq n)$ suponiendo

$$1 \leq i_1 < i_2 \ldots < i_m \leq n, \qquad 1 \leq j_1 < j_2 \ldots < j_m \leq n$$

y por I′ y J′ los complementarios respectivos de I y J en relación a $[1, n]$, estando los elementos de estos complementarios colocados en el orden natural.
Si A es una matriz cuadrada de orden n, $A = (\alpha_i^j)$ se pone $D = \det A$ y se designa por D_I^J el determinante de la submatriz de A obtenida suprimiendo en A las columnas de índice $i \in I'$ y de índice $j \in J'$.
Se pone, en fin

$$S_I^J = \sum_{\substack{p \in S_n \\ i \in I, p(i) \in J}} \varepsilon(p)\alpha_1^{p(1)} \ldots \alpha_n^{p(n)}.$$

a) Se toma $I = \{1, 2, \ldots, m\}$, demostrar que

$$S_I^I = D_I^I D_{I'}^{I'}$$

(utilizar el procedimiento del ejercicio 262);
b) I y J son dos partes cualesquiera de $[1, n]$ con m elementos, demostrar que

$$S_I^J = (-1)^{i_1 + \ldots + i_m + j_1 + \ldots + j_m} D_I^J D_{I'}^{J'}$$

(utilizar un procedimiento análogo al del § 170 para calcular el adjunto β_j^i de α_i^j):
c) Dejando $I = \{i_1, \ldots, i_m\}$ *fijo*, se supone que J describe el conjunto $\mathcal{J}$ de las partes de m elementos de $[1, n]$ (supuesto escritas *en el orden natural*), *demostrar* la fórmula de LAPLACE).

$$(1) \qquad D = (-1)^{i_1 + \ldots + i_m} \sum_{J \in \mathcal{J}} (-1)^{j_1 + \ldots + j_m} D_I^J D_{I'}^{J'}.$$

Hallar una fórmula (2) análoga a (1) dejando J fijo y I variable.
d) Demostrar que si I es fijo y si K es una parte de $[1, n]$ de m elementos distinta de I′ se tiene

$$(1') \qquad \sum_{J \in \mathcal{J}} (-1)^{j_1 + \ldots + j_m} D_I^J D_K^{J'} = 0.$$

Calcular los determinantes siguientes con la regla de LAPLACE (V. ejercicio precedente).

265.

$$\begin{vmatrix} 0 & a & b & c \\ -a & 0 & r & -q \\ -b & -r & 0 & p \\ -c & q & -p & 0 \end{vmatrix} \qquad \text{(tomar } m = 2\text{).}$$

266.

$$\begin{vmatrix} x & py & qz & -pqt \\ y & x & qt & -qz \\ z & -pt & x & py \\ t & -z & y & x \end{vmatrix} \qquad \text{(tomar } m = 2\text{).}$$

En los ejercicios siguientes (267 al 272) *se podrá o bien utilizar el método del* § 173, *o bien considerar a* A *como asociada a* f *(las columnas de* A *siendo* $f(a_i) = b_i$ $(1 \leq i \leq n)$ *se buscará expresar* a_i *mediante* $b_1, \ldots, b_n$).

267. Siendo t real hallar la inversa de

$$\begin{pmatrix} \cos t & -\operatorname{sen} t \\ \operatorname{sen} t & \cos t \end{pmatrix} \qquad \begin{pmatrix} \operatorname{ch} t & \operatorname{sh} t \\ \operatorname{sh} t & \operatorname{ch} t \end{pmatrix} \qquad \begin{pmatrix} \operatorname{sh} t & \operatorname{ch} t \\ \operatorname{ch} t & \operatorname{sh} t \end{pmatrix}.$$

268. Hallar las inversas de las matrices

$$\begin{pmatrix} -3 & 2 & -1 \\ 2 & 0 & 1 \\ -1 & 2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 2 & 2 & -1 \\ -2 & 1 & 4 \\ 3 & 0 & 0 \end{pmatrix}.$$

269. Hallar las inversas de las matrices ($i^2 = -1$; $j^3 = 1$, $j \neq 1$)

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & i\sqrt{2} & -i\sqrt{2} \\ 1 & -1 & -1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 1 \\ 1 & j & j^2 \\ 1 & j^2 & j \end{pmatrix}.$$

270. Hallar la inversa de la matriz $A = (\alpha_i^j)$ de orden n tal que

$$i \geq j \quad \alpha_i^j = 1; \qquad i < j \quad \alpha_i^j = 0.$$

271. Invertir las matrices A_n siguientes (orden n) cuando sean inversibles

$$A_3 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \qquad A_4 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \qquad A_n = \begin{pmatrix} 1 & 1 & 0 \dots 0 & 0 \\ 0 & 1 & 1 \dots 0 & 0 \\ & & & \\ 0 & 0 & 0 \dots 1 & 1 \\ 1 & 0 & 0 \dots 0 & 1 \end{pmatrix}$$

272. Invertir una matriz triangular de orden n en la que todos los elementos diagonales son inversibles.

273. Tenemos la siguiente matriz de elementos en $A = \mathbf{Z}/n\mathbf{Z}$

$$\begin{pmatrix} \dot{1} & \dot{2} & \dot{3} \\ \dot{2} & \dot{0} & \dot{1} \\ \dot{1} & \dot{2} & \dot{1} \end{pmatrix}$$

Es inversible en $\mathbf{M}_3(A)$, **a)** para $n = 5$, **b)** para $n = 9$.

274. Hallar el rango de las matrices siguientes:

$$\begin{pmatrix} 1 & 7 & 5 & 3 & -2 \\ 0 & 4 & 2 & 2 & 0 \\ 2 & -2 & 4 & 0 & 1 \\ 3 & -1 & 7 & 1 & 3 \end{pmatrix} \qquad \begin{pmatrix} 2 & -3 & 4 \\ 3 & 1 & 5 \\ -1 & 0 & -1 \\ 0 & 2 & 4 \end{pmatrix}$$

275. ¿Cuál es el rango de la matriz definida en el ejercicio n.º 237
a) Cuando $\mathbf{K} = \mathbf{R}$, **b)** cuando $\mathbf{K} = \mathbf{C}$?

276. ¿Cuál es el rango de la matriz A_n del ejercicio n.º 271?

10

ECUACIONES LINEALES

Las distintas partes de este capítulo hubieran podido situarse después de algunos párrafos de los capítulos 7, 8 y 9 de los que son aplicaciones directas. Hemos pensado, dado la importancia teórica y práctica de las ecuaciones lineales, que sería útil para el lector encontrar agrupadas en el mismo capítulo las definiciones y los resultados relativos a estas ecuaciones.

Salvo indicación contraria, las matrices y los determinantes considerados en este capítulo tienen sus elementos en un cuerpo conmutativo K.

175. Definición y propiedades generales

a) Ecuación lineal

Dados dos espacios vectoriales E y F sobre el mismo cuerpo conmutativo K, a toda aplicación lineal f de E en F y a todo elemento b de F se puede asociar una ecuación (§ 14, *b*)

$$f(x) = b \tag{1}$$

llamada *ecuación lineal*: b se llama el *segundo miembro*, si $b = 0$, la ecuación $f(x) = 0$ se llama *ecuación lineal y homogénea* asociada a la ecuación lineal (1).

Resolver (1) es hallar todos los elementos x de $f^{-1}(b) \subset E$, llamados *soluciones* de la ecuación (1). Si $f^{-1}(b) = \varnothing$ se dice que la ecuación (1) es *imposible*. Las soluciones de $f(x) = 0$ son los elementos de Ker f; ahora bien, el núcleo de f contiene siempre 0, luego una ecuación lineal y homogénea no es nunca imposible, admite siempre la solución $x = 0$, llamada *solución trivial*.

La discusión presentada en el § 14, *b*) es válida para la ecuación lineal $f(x) = b$ (como para toda ecuación asociada a una aplicación cualquiera). Pero la linealidad proporciona resultados más precisos:

Sea x_0 una solución particular de (1), se tiene $f(x_0) = b$, de donde $f(x - x_0) = 0$; luego $z = x - x_0$ es una solución de $f(x) = 0$; recíprocamente, $f(z) = 0$ implica $f(x_0 + z) = f(x_0) + f(z) = b$, luego:

TEOREMA 1. — *Si x_0 es una solución de la ecuación lineal $f(x) = b$, el conjunto de las soluciones de esta ecuación está constituido por los elementos $x_0 + z$, donde z describe* Ker f.

Teniendo en cuenta los resultados de la discusión recordada anteriormente y las propiedades de f (§ 140, corolario del teorema 2) se obtiene:

Corolario. — *Para la ecuación lineal $f(x) = b$, las propiedades siguientes son equivalentes:*

1. *f es inyectivo.*
2. Ker $f = \{0\}$.
3. *La ecuación $f(x) = b$ tiene a lo sumo una solución para todo b.*
4. *La ecuación $f(x) = 0$ sólo tiene la solución trivial.*

Para resolver una ecuación lineal es suficiente, pues, encontrar el núcleo de f y una solución particular x_0.

OBSERVACION

La linealidad de f implica el resultado siguiente: si $b = b_1 + \dots + b_i + \dots + b_n$ y si x_i es una solución de $f(x) = b_i$ $(1 \leq i \leq n)$, $x_1 + \dots + x_i + \dots + x_n$ es una solución de $f(x) = b$.

b) Sistema de ecuaciones lineales

Sea E un espacio vectorial sobre K, (F_i) una familia finita de espacios vectoriales sobre K, cuyos índices son $I = [1, n]$, y para todo i de I, una aplicación lineal f_i de E en F_i, toda familia de ecuaciones lineales

$$\text{(S)} \qquad (i = 1, 2, \dots, n) \qquad f_i(x) = b_i \qquad (b_i \in F_i)$$

se llama *sistema de ecuaciones lineales.*

Todo sistema de ecuaciones lineales es equivalente a una sola ecuación lineal: sea F el espacio vectorial producto $F_1 \times F_2 \times \dots \times F_n$, pongamos $b = (b_1, b_2, \dots, b_n) \in F$ y para todo x de E

$$f(x) = [f_1(x), f_2(x), \dots, f_n(x)] \in F$$

f es una aplicación lineal de E en F tal que $f_i = pr_i \circ f$; se ve inmediatamente que el sistema (S) es equivalente a la única ecuación lineal

$$\text{(E)} \qquad f(x) = b.$$

Un caso particularmente importante es aquel en que $F_1 = \dots = F_n = K$; f_i es entonces una *forma lineal* definida sobre E, la designaremos l^{*i} para recordar que es un elemento del dual E* de E; b es un elemento de K^n, la designaremos $b = (b^1, \dots, b^i, \dots, b^n)$, donde b^i es entonces la i-ésima coordenada de b (respecto a la base canónica de K^n).

El sistema (S) se escribe entonces (ver § 149)

$$(i = 1, 2, \dots, n) \qquad \langle x, l^{*i} \rangle = b^i$$

se dice que cada una de estas ecuaciones es una *ecuación escalar* y que (S) es un *sistema escalar.* Es lo mismo estudiar un sistema de n ecuaciones escalares o una ecuación única $f(x) = b$: *se preferirá en general la ecuación única*

en los estudios teóricos (existencia y unicidad de las soluciones, por ejemplo) *y el sistema escalar en la determinación efectiva de las soluciones.*

EJEMPLOS Y EJERCICIOS

1. $E = F = K$, $ax = b$.
$a \neq 0$ solución única $x = a^{-1}b$.
$a = 0$, $b \neq 0$ ninguna solución.
$a = 0$, $b = 0$ todo x de K es solución.

Resolver $\dot{7}\dot{x} = \dot{4}$, en $K = \mathbf{Z}/11\mathbf{Z}$.

2. Si $E = F = K^2$: $ax + by = c$, $a'x + b'y = c'$.
La discusión y la resolución es la misma que en **R**.

Resolver $\dot{3}\dot{x} + \dot{7}\dot{y} = \dot{8}$, $\dot{2}\dot{x} + \dot{6}\dot{y} = \dot{1}$ en $K = \mathbf{Z}/13\mathbf{Z}$.

Discutir $\dot{3}\dot{x} + \dot{7}\dot{y} = \dot{8}$, $\dot{2}\dot{x} + \dot{a}\dot{y} = \dot{b}$ en $K = \mathbf{Z}/13\mathbf{Z}$.

3. El resto del capítulo se va a dedicar al estudio de la ecuación $f(x) = b$, f aplicación lineal de K^m en K^n.

4. Si $a_0, \ldots, a_n$, b son $n + 1$ funciones reales dadas de la variable real t y x una función real de t, n veces derivable, incógnita, la ecuación

$$a_0 x^{(n)} + \ldots + a_{n-1}x' + a_n x = b$$

es una ecuación lineal llamada *ecuación diferencial lineal de orden n* (ver curso de Análisis).

5. Siendo b un número real dado y x una función real continua sobre $[0, 1]$, desconocida, $\int_0^1 x(t)\,dt = b$ es una ecuación lineal.

De una manera más general dada una función real Φ definida por $(t, \theta) \to \Phi(t, \theta)$, continua para $0 \leq t \leq 1$, $0 \leq \theta \leq 1$ y una función real b continua sobre $[0, 1]$, la ecuación en que la incógnita es una función real x, continua sobre $[0, 1]$

$$\int_0^1 \Phi(t, \theta)x(\theta)\ d\theta = b(t)$$

es una ecuación lineal.

El estudio de tales ecuaciones llamadas ecuaciones integrales no se trata en matemáticas generales.

176. Sistemas de n ecuaciones de m incógnitas sobre un cuerpo conmutativo K

a) Definiciones y notaciones

En el último caso estudiado F es de dimensión finita n sobre K, supongamos ahora, además, que E sea de dimensión finita m, luego $\dim E = m$ y $\dim F = n$.

Respecto a *bases especificadas* en E y F

(1) $$f(x) = b$$

será equivalente al sistema

(2) $$\sum_{i=1}^{m} a_i^j x^i = b^j \qquad (j = 1, 2, \ldots, n)$$

que se puede escribir de manera desarrollada

(2) $$\begin{cases} a_1^1 x^1 + \ldots + a_i^1 x^i + \ldots + a_m^1 x^m = b^1 \\ \vdots \\ a_1^j x^1 + \ldots + a_i^j x^i + \ldots + a_m^j x^m = b^j \\ \vdots \\ a_1^n x^1 + \ldots + a_i^n x^i + \ldots + a_m^n x^m = b^n \end{cases}$$

Recíprocamente todo sistema de esta forma puede ser asociado a una y sólo una aplicación lineal f de K^m en K^n (expresados en sus bases canónicas) y al elemento $b = (b^1, \ldots, b^n)$ de K^n.

Poniendo $A = (a_i^j)$ $(1 \leq i \leq m, 1 \leq j \leq n)$ el sistema (2) puede también escribirse

(2') $$A \begin{pmatrix} x^1 \\ \vdots \\ x^i \\ \vdots \\ x^m \end{pmatrix} = \begin{pmatrix} b^1 \\ \vdots \\ b^i \\ \vdots \\ b^n \end{pmatrix}.$$

Consideremos los vectores filas de A, l^{*j} $(1 \leq j \leq n)$, elemento de $(K^m)^*$ (dual del espacio de salida K^m de f); $a_1^j, \ldots, a_i^j, \ldots, a_m^j$ son las coordenadas de l^{*j} en relación a la base de $(K^m)^*$, dual de la base canónica de K^m, luego

$$\sum_{i=1}^{m} a_i^j x^i = \langle x, l^{*j} \rangle \qquad (j = 1, 2, \ldots, n)$$

el sistema (2) puede también escribirse

(2'') $$\begin{cases} \langle x, l^{*1} \rangle = b^1 \\ \vdots \\ \langle x, l^{*j} \rangle = b^j \\ \vdots \\ \langle x, l^{*n} \rangle = b^n \end{cases}$$

finalmente considerando los vectores columnas c_i $(1 \leq i \leq m)$ de A (elementos de K^n espacio de llegada de f); $a_i^1, \ldots, a_i^j, \ldots, a_i^n$ son las coordenadas de c_i respecto a la base canónica de K^n, (2) puede también escribirse

(3) $$x^1 c_1 + \ldots + x^i c_i + \ldots + x^m c_m = b$$

y la resolución de (1) o (2) consiste en escribir efectivamente el vector b de K^n como combinación lineal de m vectores $c_1, \ldots, c_i, \ldots, c_m$ de K^n.

Los escalares $x^1, \ldots, x^i, \ldots, x^m$ se llaman las *incógnitas* del sistema escalar (2'): se dice entonces que (2') es un *sistema de n ecuaciones (escalares) con m incógnitas*: los elementos del cuerpo conmutativo K, a_i^j y b^j se llaman, respectivamente, los *coeficientes de las incógnitas* y los *segundos miembros*.

El sistema

$$\sum_{i=1}^{m} a_i^j x^i = 0 \qquad (j = 1, 2, \ldots, n)$$

se llama *sistema lineal y homogéneo asociado al sistema*

$$\sum_{i=1}^{m} a_i^j x^i = b^j \qquad (j = 1, 2, \ldots, n).$$

La matriz A se llama *matriz del sistema*.

Se llama *rango r* del sistema (2') el rango de la matriz A (o de la aplicación lineal f): es el rango de los vectores columnas c_i y de los vectores filas l^{*j}.

Es también el orden de un menor de A de orden máximo no nulo (ver § 174, b).

Habiendo *elegido* un tal menor Δ de orden r (suponiendo que exista) se dice que Δ es el *determinante principal* del sistema. Las incógnitas cuyos coeficientes figuran (resp. no figuran) en Δ se llaman *incógnitas principales* (resp. no principales). Las ecuaciones en las que ciertos coeficientes figuran (resp. no figuran) en Δ se llaman *ecuaciones principales* (resp. no principales). El sistema de las ecuaciones principales se llama *sistema principal* del sistema (2).

Observemos también que todas estas denominaciones no son intrínsecas, sino relativas a la elección del menor Δ.

Tenemos, pues, varios medios de estudiar un sistema escalar de n ecuaciones escalares con m incógnitas escalares:

— Estudio *intrínseco* (válido en casos mucho más generales: teorema 1 y corolario del § 175).

— Estudio con la ayuda de los vectores *filas* (formas lineales l^{*j}).

— Estudio con la ayuda de los vectores *columnas* (vectores c_i).

— Finalmente veremos que el empleo de los *determinantes* permite resolver y discutir todo sistema escalar.

b) Sistemas de Cramer

Un caso particularmente simple es aquel en que f es un isomorfismo de E y de F, luego $r = \text{rg } f = \dim E = \dim F = n$; A es entonces una matriz cuadrada de orden n y de rango n, y para todo b

$$f(x) = b \Leftrightarrow x = f^{-1}(b)$$

de donde utilizando los resultados enunciados en el § 169, b):

Teorema 2 y definición. — *Si* E *y* F *son dos espacios de dimensión* n *sobre un cuerpo conmutativo* K, f *una aplicación lineal de* E *en* F, $A = M(f)$, *las propiedades siguientes son equivalentes:*

1. f *es un isomorfismo.*
2. $\operatorname{rg}(f) = \operatorname{rg}(A) = n$.
3. *Los vectores columna* (c_i) *y los vectores fila* (l^{*j}) *de* A *son independientes.*
4. A *es inversible.*
5. $\det A \neq 0$.
6. *Para todo* b *de* F, $f(x) = b$ *tiene una única solución.*

Se dice que el sistema

$$\sum_{n}^{i=1} a_i^j x^i = b^j \qquad (j = 1, 2, \ldots, n)$$

es un sistema de Cramer.

Corolario. — *Un sistema escalar de* n *ecuaciones, con* n *incógnitas es un sistema de Cramer si y sólo si el sistema homogéneo asociado no admite más que la solución trivial* $(0, \ldots, 0, \ldots, 0)$.

Un caso particularmente sencillo de sistema de Cramer es el siguiente

$$\begin{cases} a_1^1 x^1 & = b^1 \\ \vdots & \\ a_1^n x^1 + \ldots + a_i^n x^i + \ldots + a_n^n x^n & = b^n \\ \vdots & \\ a_1^i x^1 + \ldots + a_i^i x^i & = b^i \end{cases}$$

en el caso en que *para todo* i de $[1, n]$ $a_i^i \neq 0$. En efecto, la matriz A es triangular inferior y todos los elementos de la diagonal principal son no nulos, luego (§ 170, *b*)

$$\det A = a_1^1 \ldots a_i^i \ldots a_n^n \neq 0.$$

Se calcula sin ambigüedad x^1, después x^2, ..., a continuación x^n. A un tal sistema le llamaremos *sistema de Cramer triangular.*

OBSERVACION

La fórmula (2′) del § 176 da para un sistema de Cramer

$$\begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = A^{-1} \begin{pmatrix} b^1 \\ \vdots \\ b^n \end{pmatrix}.$$

El cálculo efectivo de esta solución por uno de los métodos que vamos a exponer en los párrafos siguientes (177 al 179) da

$$(j = 1, \ldots, n) \qquad x^j = \sum_{i=1}^{n} \lambda_i^j b^i$$

y $A^{-1} = (\lambda_i^j)$: este método de calcular A^{-1} es algunas veces más simple que el indicado en el § 173.

177. Estudio con ayuda de los vectores columna

Dados los vectores $c_1, \ldots, c_i, \ldots, c_m$ y b de K^n buscamos los escalares $x^1, \ldots, x^i, \ldots, x^m$ tales que

$$(3) \qquad x^1c_1 + \ldots + x^ic_i + \ldots + x^mc_m = b.$$

Es necesario primero determinar el rango r de los vectores $c_1, \ldots, c_m$, se nos ha enseñado cómo hay que hacerlo en el § 138 sea r este rango, ($r \leq \inf(m, n)$).

a) Supongamos $m = n = r$. El sistema correspondiente es un sistema de CRAMER. Apliquemos el procedimiento del § 138 a los vectores $c_1, \ldots, c_m, b$; ello nos dice que son ligados. Encontramos, pues, una relación de dependencia que nos dará $x^1, \ldots, x^i, \ldots, x^m$.

EJERCICIO

1. Sea el sistema (ver § 138, ej. 2)

$$\begin{cases} 2x - y + 4z = -4 \\ 3x + 2y - 3z = 17 \\ 5x - 3y + 8z = -10 \end{cases}$$

los vectores c_1, c_2, c_3 son, respectivamente, los vectores x_1, x_2, x_3 y b el vector x_4 tratado en el ejemplo citado. Hemos encontrado que c_1, c_2, c_3 eran independientes, de modo que el sistema es un sistema de CRAMER, y

$$b = 2c_1 + 4c_2 - c_3$$

luego

$$x = 2, \; y = 4, \; z = -1.$$

b) Supongamos $r = n < m$. *No hay ecuaciones no principales,* pero *hay incógnitas no principales*; supongamos, mediante un cambio de numeración de incógnitas, que las r primeras incógnitas son principales. Demos a las $m - n$ incógnitas no principales valores arbitrarios de K: $x^{n+1}, \ldots, x^m$, la ecuación (3) se escribe

$$(3') \qquad x^1c_1 + \ldots + x^nc_n = b - (x^{n+1}c_{n+1} + \ldots + x^mc_m) = b'$$

$c_1, c_2, \ldots, c_n$ es una base de K^n; $c_{n+1}, \ldots, c_m$ pertenecen a K^n igual que b, luego b' también pertenece: podremos, en consecuencia, como más arriba, descomponer b' sobre $c_1, \ldots, c_n$ y calcular efectivamente $x^1, \ldots, x^n$, naturalmente estos valores dependerán de los escalares arbitrarios $x^{n+1}, \ldots, x^m$; diremos que existe una *indeterminación* de orden $m - n$.

Si A_n es la matriz del sistema principal, $\det A_n \neq 0$, se tendrá

$$\begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} = (A_n)^{-1} \begin{pmatrix} b^1 - \sum_{k=n+1}^{m} x^k a_k^1 \\ \vdots \\ b^n - \sum_{k=n+1}^{m} x^k a_k^n \end{pmatrix}$$

se ve, por lo tanto, que $x^1, \ldots, x^n$ se expresan linealmente en función de $b^1, \ldots, b^n, x^{n+1}, \ldots, x^m$.

c) Si $r < n$ hay *ecuaciones no principales* y eventualmente incógnitas no principales. Por cambio de numeración de las ecuaciones y temporalmente de las incógnitas supondremos que las ecuaciones e incógnitas principales son las r primeras; demos valores arbitrarios $x^{r+1}, \ldots, x^m$ a las incógnitas no principales, si existen, y escribamos la ecuación (3) bajo la forma

$$(3'') \qquad x^1c_1 + \ldots + x^rc_r = b - (x^{r+1}c_{r+1} + \ldots + x^mc_m) = b''$$

pero en este caso $c_1, c_2, \ldots, c_r$ engendra un subespacio V de dimensiones $r < n$ de K^n, (3″) no tiene solución más que si $b'' \in V$. Pero cualquiera que sean los valores $x^{r+1}, \ldots, x^m$, el vector $x^{r+1}c_{r+1} + \ldots + x^mc_m \in V$, para que (3), luego (3″), tenga solución es necesario y suficiente que $b \in V$; ahora bien, $b \in K^n$ y V es un subespacio de dimensión $r < n$ de K^n: para que (3″) tenga solución habrá que añadir en general $n - r$ condiciones escalares: si se completa la base $c_1, \ldots, c_r$ de V con $d_{r+1}, \ldots, d_n$ para obtener una base de K^n, habrá que expresar que las $n - r$ últimas coordenadas de b en la base $c_1, \ldots, c_r, d_{r+1}, \ldots, d_n$ son nulas.

d) Vemos, pues, que el procedimiento de discusión y resolución del sistema escalar (2) con ayuda de la ecuación (3) no es práctico más que si $r = n$, es decir, si no hay ecuaciones no principales. Si $r < n$ el procedimiento nos conduce, de hecho, a cambiar de base en K^n: pasando de la base canónica a la base $c_1, \ldots, c_r, d_{r+1}, \ldots, d_n$ lo que muchas veces resulta complicado en la práctica.

Se podría también pensar en resolver el sistema principal y al imponer que las soluciones de este sistema verifiquen las ecuaciones no principales se constata que las condiciones obtenidas son independientes de los valores de $x^{r+1}, \ldots, x^m$ atribuidos a las incógnitas no principales (si existen): resulta evidente, pues $b - b'' \in V$. Este método sólo es fácil en los casos simples (ver ejercicios 2 y 3 más abajo); en el caso general el empleo de los vectores filas (ver § 178) o el empleo de los determinantes (ver § 179) da más fácilmente las condiciones de existencia de las soluciones.

EJERCICIOS

2. $K = \mathbf{R}$ discutir el sistema

$$cy - bz = l, \qquad az - cx = m, \qquad bx - ay = n$$

(se demostrará que $r \leq 2$ y que $r = 2$ si $a^2 + b^2 + c^2 \neq 0$; suponiendo $r \neq 0$ se resolverá en x, y las dos primeras ecuaciones y se mirará si esta solución del sistema principal verifica la tercera ecuación).

3. $K = \mathbf{R}$ discutir el sistema

$$\lambda x + y + z = a, \qquad x + \lambda y + z = b, \qquad x + y + \lambda z = c$$

utilizando los vectores columnas.

(Se obtiene una resolución y una discusión mucho más rápida introduciendo la incógnita auxiliar $s = x + y + z$).

178. Estudio mediante los vectores filas (formas lineales)

Consideremos el sistema de n ecuaciones escalares con m incógnitas escrito en la forma

$$(2'') \quad \begin{cases} \langle x, l^{*1} \rangle = b^1 \\ \vdots \\ \langle x, l^{*i} \rangle = b^i \\ \vdots \\ \langle x, l^{*n} \rangle = b^n \end{cases}$$

$l^{*i}, \ldots, l^{*n}$ son formas definidas sobre $E = K^m$.

Sea r el rango del sistema, luego el rango de estas formas (§ 162, *a*), y § 176, *a*), podemos determinarlo por un procedimiento análogo al del § 138 (operando en E*)

$$r \leq \inf (m, n).$$

a) Si $r = n$. *No hay ecuaciones no principales.* Si hay incógnitas no principales ($r = n < m$), por un cambio de numeración de las incógnitas podemos suponer que éstas son $x^{n+1}, \ldots, x^m$ (siendo $x^1, x^2, \ldots, x^n$ las incógnitas principales); podríamos resolver (2'') como lo hemos hecho en el § 177, *b*. Continuemos más bien razonando sobre las formas o mejor sobre las ecuaciones: en efecto, el procedimiento del § 138 aplicado en E* consiste en reemplazar en el sistema $\{l^{*1}, \ldots, l^{*i}, \ldots, l^{*n}\}$, l^{*j} por $\sum_{i=1}^{n} \lambda_i l^{*i}$, y los dos sistemas de formas tienen el mismo rango si $\lambda_j \neq 0$. Generalicemos este procedimiento a las *ecuaciones* del sistema (2''), si $\lambda_j \neq 0$ (2'') es equivalente al sistema de ecuaciones obtenido reemplazando la ecuación

$$\langle x, l^{*j} \rangle = b^j \qquad \text{por} \qquad \left\langle x \sum_{i=1}^{n} \lambda_i l^{*i} \right\rangle = \sum_{i=1}^{n} \lambda_i b^i.$$

Por aplicación repetida de este procedimiento obtendremos cambiando según la necesidad la numeración de las incógnitas principales en el sistema (2''') equivalente a (2'')

$$(2'')\left\{\begin{array}{l}
a'^1_1 x^1 + \ldots\ldots\ldots\ldots\ldots + a'^1_n x^n + a'^1_{n+1} x^{n+1} + \ldots + a'^1_m = b'^1 \\
\quad\ddots \\
\qquad a'^2_2 x^2 + \ldots\ldots\ldots\ldots + a'^2_n x^n + a'^2_{n+1} x^{n+1} + \ldots + a'^2_m = b'^2 \\
\qquad\quad\ddots \\
\qquad\qquad a'^j_j x^j + \ldots + a'^j_n x^n + a'^j_{n+1} x^{n+1} + \ldots + a'^j_m = b'^j \\
\qquad\qquad\quad\ddots \\
\qquad\qquad\qquad a'^n_n x^n + a'^n_{n+1} x^{n+1} + \ldots + a'^n_m = b'^n
\end{array}\right.$$

para todo j de $[1, n]$, $a'^j_j \neq 0$.

Es decir, obtenemos con relación a las *incógnitas principales un sistema de Cramer triangular.* Encontramos de nuevo con ello el procedimiento elemental de resolución de las ecuaciones lineales conocido con el nombre de "método de adición".

Se da a las incógnitas no principales (si existen) valores $x^{n+1}, \ldots, x^m$ arbitrarios y se calcula sucesivamente $x^n, x^{n-1}, \ldots, x^1$.

El sistema tiene, pues, las soluciones con una indeterminación de orden $m - n$.

EJERCICIOS

1. Escribiendo para simplificar, f por $f(x, y, z)$ ... obtenemos sucesivamente los sistemas equivalentes

$$\left\{\begin{array}{l}
f = 2x - y + 4z = -4 \\
g = 3x + 2y - 3z = 17 \\
h = 5x - 3y + 8z = -10
\end{array}\right.$$

$$\left\{\begin{array}{l}
f' = f = 2x - y + 4z = -4 \\
g' = 3f - 2g = -7y + 18z = -46 \\
h' = 5f - 2g = y + 4z = 0
\end{array}\right.$$

$$\left\{\begin{array}{l}
f'' = f' = 2x - y + 4z = -4 \\
g'' = g' = -7y + 18z = -46 \\
h'' = g' + 7h' = 46z = -46.
\end{array}\right.$$

De donde $z = -1$, $y = 4$, $x = 2$.

2. Resolver

$$\begin{cases} 2x - y + 4z + t = -4 \\ 3x + 2y - 3z - 5t = 17 \\ 5x - 3y + 8z + 2t = -10. \end{cases}$$

b) Si $r < n$, *hay ecuaciones no principales*; por un cambio de numeración podemos suponer que son las r primeras las que constituyen el sistema principal. Las formas $l^{*1}, \ldots, l^{*r}$ son independientes y existen escalares λ_k^j tales que

$$(4) \qquad l^{*j} = \lambda_1^j l^{*1} + \ldots + \lambda_r^j l^{*r} \qquad (j = r+1, \ldots, n).$$

Si el sistema (2″) tiene una solución $x = (x^1, \ldots, x^m)$ tendremos

$$(4') \qquad b^j = \lambda_1^j b^1 + \ldots + \lambda_r^j b^r \qquad (j = r+1, \ldots, n).$$

Recíprocamente sea x una solución del sistema principal, está claro que si las $n - r$ relaciones (4′) se satisfacen x será una solución del sistema (2″), de donde:

TEOREMA 3. — *Dado un sistema de n ecuaciones escalares con m incógnitas*

$$(2'') \qquad \langle x, l^{*j} \rangle = b^j \qquad (j = 1, 2, \ldots, n)$$

las formas $l^{*1}, \ldots, l^{*r}$ *siendo independientes y si las formas* $l^{*r+1}, \ldots, l^{*n}$ *verifican*

$$(4) \qquad l^{*j} = \lambda_1^j l^{*1} + \ldots + \lambda_r^j l^{*r} \qquad (j = r+1, \ldots, n)$$

toda solución del sistema principal (sistema de las r primeras ecuaciones) es solución del sistema (2″) *si y sólo si*

$$(4') \qquad b^j = \lambda_1^j b^1 + \ldots + \lambda_r^j b^r \qquad (j = r+1, \ldots, n).$$

Las λ_k^j se obtienen aplicando al sistema de formas (l^{*j}) $(1 \leq j \leq n)$ el método del § 138 (en $E^* = (K^m)^*$) obtenemos de una manera muy simple las $n - r$ relaciones (4′) que se llaman las *condiciones de compatibilidad* del sistema considerado.

Si estas condiciones no se satisfacen el sistema es imposible; si se satisfacen, las soluciones del sistema son las del sistema principal: ha sido estudiado más arriba en *a*), habrá una indeterminación de orden $m - r$.

EJERCICIO

3. Discutir por este método el sistema del ejercicio 2 del § 177.

179. Estudio mediante los determinantes

El rango del sistema es el orden máximo de un menor extraído de la matriz A. Modificando eventualmente la numeración de las incógnitas y el de las ecuaciones supondremos que

$$\Delta = \begin{vmatrix} a_1^1 \ldots a_r^1 \\ \vdots \quad\quad \vdots \\ a_1^r \ldots a_r^r \end{vmatrix} \neq 0.$$

a) $m = n = r$. Tenemos un sistema de CRAMER, hay una solución única que podemos encontrar del modo siguiente: sean $c_1, c_2, \ldots, c_n$ los vectores columna de la matriz cuadrada A: son independientes. Como b, pertenecen a $F = K^n$, la solución $x^1, \ldots, x^n$ es tal que

$$(3) \qquad x^1 c_1 + \ldots + x^n c_n = b$$

de donde utilizando la propiedad 4 del § 168, *b*) tenemos
$\det(c_1, \ldots, c_{i-1}, b, c_{i+1}, \ldots, c_n) = x^i \det(c_1, \ldots, c_{i-1}, c_i, c_{i+1}, \ldots, c_n) = x^i \det A$
ahora bien, $\det A \neq 0$; de donde:

TEOREMA 4.—*Un sistema de Cramer de n ecuaciones con n incógnitas tiene por solución única*

$$x^i = (\det B_i)(\det A)^{-1} \qquad (i = 1, 2, \ldots, n)$$

donde A *es la matriz* (a_i^j) *de los coeficientes de las incógnitas y* B_i *la matriz obtenida reemplazando en* A *el vector columna* c_i *por el vector columna* b *de los segundos miembros.*

b) Si $r = n < m$, como en el § 177, *b*) se da a las incógnitas no principales los valores arbitrarios $x^{n+1}, \ldots, x^m$ y se resuelve el sistema que, con relación a las incógnitas principales, es un sistema de CRAMER, se obtiene

$$x^i = \frac{\det\left(c_1, \ldots, c_{i-1}, b - \sum_{k=n+1}^{m} x^k c_k, c_{i+1}, \ldots, c_n\right)}{\det(c_1, \ldots, c_i, \ldots, c_n)} \qquad (i = 1, \ldots, n)$$

de donde desarrollando el determinante numerador respecto a la columna se obtiene los escalares λ_i^i, μ_k^i tales que

$$x^i = \lambda_i^i b^i + \mu_{n+1}^i x^{n+1} + \ldots + \mu_m^i x^m \qquad (i = 1, 2, \ldots, n)$$

los escalares λ_i^i $(1 \leq i \leq n)$, así como los escalares μ_k^i $(n + 1 \leq k \leq m$, $1 \leq i \leq n)$, no dependen más que de los coeficientes a_i^j pertenecientes a la matriz del determinante principal.

c) Si $r < n$, hay ecuaciones no principales.

$c_1, c_2, \ldots, c_r$ son independientes. Como en el § 177, *c*) demos a las incógnitas no principales (si existen) valores arbitrarios $x^{r+1}, \ldots, x^m$, tendremos

$$(3'') \qquad x^1 c_1 + \ldots + x^r c_r = b - (x^{r+1} c_{r+1} + \ldots + x^m c_m) = b''$$

para que b'' pertenezca al espacio V engendrado por $c_1, \ldots, c_r$ es necesario y suficiente que b pertenezca a V (§ 177, *c*): hemos tratado este problema en el § 174, siendo el resultado dado por el lema, de donde:

TEOREMA 5.—*Dado un sistema de n ecuaciones escalares con m incógnitas*

$$(2) \qquad \sum_{i=1}^{m} a_i^j x^i = b^j \qquad (j = 1, 2, \ldots, n)$$

de rango r, siendo independientes los vectores columnas $c_1, \ldots, c_r$ de la matriz de los coeficientes, toda solución del sistema principal es solución del sistema (2), si y sólo si los $n - r$ determinantes característicos de b son nulos.

Si uno al menos de estos determinantes característicos es no nulo, el sistema es imposible.

EJERCICIOS

1. Estudiar y resolver eventualmente todos los sistemas propuestos en los ejemplos y ejercicios de los §§ 177 y 178.

2. Sea un sistema de $n + 1$ ecuaciones con n incógnitas

$$\sum_{i=1}^{n} a_i^j x^i = b^j \qquad (j = 1, \ldots, n + 1)$$

se llama *determinante completo* del sistema el determinante $\Delta_{n+1} = \det(c_1, \ldots, c_n, b)$. Demostrar que si el sistema es de rango n, el sistema propuesto tiene solución si y sólo si $\Delta_{n+1} = 0$. ¿Qué sucederá si el rango del sistema es $r < n$?

180. Estudio particular de los sistemas escalares homogéneos

Con las notaciones del § 176, siendo f una aplicación de K^m en K^n (o de un espacio E de dimensión m en un espacio f de dimensión n) la resolución de

$$f(x) = 0 \tag{1}$$

es equivalente a la del sistema

$$\text{(2)} \qquad \begin{cases} \langle x, l^{*1} \rangle = a_1^1 x^1 + \ldots + a_m^1 x^m = 0 \\ \vdots \\ \langle x, l^{*n} \rangle = a_1^n x^1 + \ldots + a_m^n x^m = 0. \end{cases}$$

Según (1), las soluciones de $f(x) = 0$ son los elementos del núcleo de f y si f es de rango r, Ker f es de dimensión $m - r$ (ver § 143, teorema 7).

Según (2), las soluciones de (1) son elementos del ortogonal en E de V^* (ver § 151), como V^* está engendrado por las formas $l^{*1}, \ldots, l^{*n}$, cuyo rango es igualmente r, resulta

$$(V^*)^{\perp} = \text{Ker } f.$$

El sistema (2) no es nunca imposible, admite siempre la solución trivial $x^1 = \ldots = x^m = 0$, sólo admite ésta si f es inyectiva, es decir, si Ker $f = \{0\}$, o sea, si $r = m$.

Si $r < m$, Ker f es no nulo: las fórmulas del § 179 nos permitirán encontrar una base de este subespacio. r es igualmente el orden máximo de un menor extraído de la matriz de los coeficientes del sistema. Cambiando si fuera preciso la numeración de las incógnitas y de las ecuaciones del sistema (2) podemos suponer que

$$\Delta = \begin{vmatrix} a_1^1 \dots a_r^1 \\ \vdots \quad\quad \vdots \\ a_1^r \dots a_r^r \end{vmatrix} \neq 0.$$

Las soluciones del sistema (2) son las soluciones del sistema principal (todas las características son nulas, al ser nulos todos los segundos miembros). Demos a $x^{r+1}, \dots, x^m$ valores arbitrarios $\lambda^{r+1}, \dots, \lambda^m$ tendremos

$$\begin{cases} x^1 = \mu_{r+1}^1 \lambda^{r+1} + \dots + \mu_m^1 \lambda^m \\ \vdots \\ x^r = \mu_{r+1}^r \lambda^{r+1} + \dots + \mu_m^r \lambda^m \\ x^{r+1} = \lambda^{r+1} \\ \vdots \\ x^m = \lambda^m. \end{cases}$$

Las μ_i^j $(r+1 \leq i \leq m,\ 1 \leq j \leq r)$ no dependen más que de a_i^j; designemos por v_i $(r+1 \leq i \leq m)$ el vector que tiene por coordenadas los coeficientes de λ^i en las fórmulas precedentes, tendremos

$$(5) \qquad x = \lambda^{r+1} v_{r+1} + \dots + \lambda^m v_m$$

siendo $(\lambda^{r+1}, \dots, \lambda^m)$ un elemento arbitrario de K^{m-r}, los vectores $v_{r+1}, \dots, v_m$ engendran, pues, Ker f, hay $m-r$, luego describen una base de Ker $f = (V^*)^{\perp}$, pues V^* está engendrado por el sistema de las n formas $l^{*1}, \dots, l^{*n}$ de rango r. Hemos dicho ya (§ 151, ej. 2) que (2) se llamaba el *sistema de ecuaciones cartesianas* de $(V^*)^{\perp}$, diremos que (5) es una *ecuación paramétrica* de este subespacio de E.

EJEMPLOS Y EJERCICIOS

1. Discutir el sistema

$$ax + by + cz = 0, \qquad a'x + b'y + c'z = 0.$$

Si este sistema es de rango 2, demostrar que (siendo λ un elemento arbitrario de K)

$$x = \lambda(bc' - cb'), \qquad y = \lambda(ca' - ac'), \qquad z = \lambda(ab' - ba').$$

2. Generalizar el sistema $(m = n + 1)$

$$(j = 1, \dots, n) \qquad a_1^j x^1 + \dots + a_{n+1}^j x^{n+1} = 0$$

cuando este sistema es de rango n demostrar que

$$\frac{x^1}{\Delta_1} = \frac{x^2}{-\Delta_2} = \dots = \frac{x^i}{(-1)^{i-1}\Delta_i} = \dots = \frac{x^{n+1}}{(-1)^n \Delta_{n+1}}$$

designando por Δ_i el determinante de la matriz obtenida suprimiendo la columna de número i en la matriz de los coeficientes del sistema.

181. Conclusión

El lector puede encontrarse desconcertado por la diversidad de métodos utilizados para resolver y discutir un sistema lineal.

Recordemos que para las cuestiones teóricas es mejor razonar sobre la aplicación lineal f. Respecto a la resolución efectiva haremos las observaciones siguientes:

Si el sistema comprende un *número pequeño* de ecuaciones de un número pequeño de incógnitas, y si los coeficientes tienen *valores numéricos específicos* (es decir, cuando no hay parámetros), el método del § 178, por "combinación lineal de las ecuaciones", es en general el más simple: conduce a un sistema de CRAMER triangular y eventualmente a condiciones independientes de las incógnitas, que indican si el sistema es posible o no. La utilización de los determinantes conduce, en general, a cálculos más complicados: habiendo escogido un determinante principal $\Delta_r \neq 0$, hay que formar $n - r$ características y seguidamente a calcular r otros determinantes para resolver el sistema principal.

Si el sistema comprende un *número pequeño* de ecuaciones y de incógnitas con *parámetros,* será conveniente acudir al "caso general", es decir, aquel en que el rango del sistema es máximo; en particular si $m = n$ este caso corresponde a $\Delta \neq 0$, Δ determinante de la matriz de los coeficientes; en este caso se podrá utilizar el mismo método que más arriba o bien emplear las formas de CRAMER. Cuando $\Delta = 0$, habrá en general interés en utilizar el método de "combinación de las ecuaciones" para estudiar separadamente cada caso correspondiente a los valores particulares de los parámetros.

Si el sistema presenta una cierta *simetría* (caso en que $m = n$, n no especificado), el determinante respeta esta simetría y puede en este caso ser el camino más cómodo.

Cuando m y n son *grandes,* y el sistema no presenta "simetría" alguna, todos los métodos expuestos en este capítulo conducen entonces a cálculos pesados y muy largos: se impone recurrir a las máquinas.

Ejercicios

Salvo indicación contraria los cálculos se harán en un cuerpo conmutativo cualquiera En la resolución de los sistemas siguientes (277 al 284) *se tomará* K = **Q**, *después se buscará las soluciones enteras.*

277.

$$\begin{cases} 3x - y + z = 5 \\ x + y - z = -2 \\ -x + 2y + z = 3 \end{cases}$$

278.

$$\begin{cases} 2x - y + z = 4 \\ -x + 3y - 5z = 1 \\ 8x - 9y + 13z = 2. \end{cases}$$

279.

$$\begin{cases} x - y + z = 3 \\ 5x + 2y - z = 5 \\ -3x - 4y + 3z = 1. \end{cases}$$

280.

$$\begin{cases} 3x + 4y + z + 2t = 3 \\ 6x + 8y + 2z + 5t = 7 \\ 9x + 12y + 3z + 10t = 13. \end{cases}$$

281.

$$\begin{cases} 3x - 5y + 2z + 4t = 2 \\ 7x - 4y + z + 3t = 5 \\ 5x + 7y - 4z - 6t = 3. \end{cases}$$

282.

$$\begin{cases} 8x + 6y + 5z + 2t = 21 \\ 3x + 3y + 2z + t = 10 \\ 4x + 2y + 3z + t = 8 \\ 7x + 4y + 5z + 2t = 18 \\ 3x + 5y + z + t = 15. \end{cases}$$

283.

$$\begin{cases} 2x + 3y + z + 2t = 4 \\ 4x + 3y + z + t = 5 \\ 5x + 11y + 3z + 2t = 2 \\ 2x + 5y + z + t = 1 \\ x - 7y - z + 2t = 7. \end{cases}$$

284.

$$\begin{cases} x + 2y + 3z + t = 3 \\ x + 4y + 5z + 2t = 2 \\ 2x + 9y + 8z + 3t = 7 \\ 3x + 7y + 7z + 2t = 12 \\ 5x + 7y + 9z + 2t = 20. \end{cases}$$

285. Sea el sistema ($\mathbf{K} = \mathbf{C}$, $j_3 = 1$, $j \neq 1$)

$$x + y + z = a, \qquad x + jy + j^2 z = b, \qquad x + j^2 y + jz = c.$$

a) Resolver el sistema; b) ¿Cómo hay que escoger a, b, c para que x, y, z sean reales?

286. Discutir según los valores de a, b, c, d el sistema

$$\begin{cases} x + 2y + 3z + 4t = a \\ 2x + 3y + 4z + t = b \\ 3x + 4y + z + 2t = c \\ 4x + y + 2z + 3t = d. \end{cases}$$

suponiendo: 1.º $\mathbf{K} = \mathbf{Q}$, 2.º $\mathbf{K} = \mathbf{Z}$, 3.º $\mathbf{K} = \mathbf{Z}/5\mathbf{Z}$.

En los ejercicios siguientes se discutirán los sistemas (287 al 298) *según los valores atribuidos a los parámetros suponiendo* $\mathbf{K} = \mathbf{C}$; *se dará eventualmente la discusión para* $\mathbf{K} = \mathbf{R}$ *si es diferente de la precedente.*

287.

$$\begin{cases} (1 - m)x + (2m + 1)y + (2m + 2)z = m \\ mx + my = 2m + 2 \\ 2x + (m + 1)y + (m - 1)z = m^2 - 2m + 9. \end{cases}$$

288.

$$\left\{\begin{array}{l}(m-2)x+2y-z=a\\ 2x+my+2z=b\\ 2mx+2(m+1)y+(m+1)z=c.\end{array}\right.$$

289.

$$\left\{\begin{array}{l}x+my+z=1\\ mx+y+(m-1)z=m\\ x+y+z=m+1.\end{array}\right.$$

290.

$$\left\{\begin{array}{l}(m-2)x+2y-z=m+2\\ 2x+my+2z=m^2+3\\ 2mx+2(m+1)y+(m+1)z=2m^3-\dfrac{m^2}{2}-\dfrac{m}{2}+5\end{array}\right.$$

291.

$$\left\{\begin{array}{l}x+y+z=1\\ ax+by+cz=d\\ a^2x+b^2y+c^2z=d^2.\end{array}\right.$$

292.

$$\left\{\begin{array}{l}x+y+z=1\\ ax+by+cz=d\\ a(a-1)x+b(b-1)y+c(c-1)z=d(d-1).\end{array}\right.$$

293.

$$\left\{\begin{array}{l}y+z+mt=a\\ z+t+mx=b\\ t+x+my=c\\ x+y+mz=d.\end{array}\right.$$

294.

$$\left\{\begin{array}{l}x+m(y+z+t)=a\\ y+m(z+t+x)=b\\ z+m(t+x+y)=c\\ t+m(x+y+z)=d.\end{array}\right.$$

295.

$$\left\{\begin{array}{l}\lambda x+\mu y+z=1\\ x+\lambda\mu y+z=\mu\\ x+\mu y+\lambda z=1.\end{array}\right.$$

296.

$$\left\{\begin{array}{l}2\lambda x+\mu y+2z=1\\ 2\lambda x+(2\mu-1)y+3z=1\\ 2\lambda x+\mu y+(\mu+3)z=2\mu-1.\end{array}\right.$$

297.

$$\left\{\begin{array}{l}x+cy+bz=-a\\ y+az+cx=-b\\ z+bx+ay=-c.\end{array}\right.$$

298.

$$\left\{\begin{array}{l}ax-by-cz-dt=p\\ bx+ay+dz-ct=q\\ cx-dy+az+bt=r\\ dx+cy-bz+at=s.\end{array}\right.$$

299. **K** = **R**

$$\left\{\begin{array}{l}x\cos 2\alpha+y\cos\alpha+z=a\\ x\cos 2\beta+y\cos\beta+z=b\\ x\cos 2\gamma+y\cos\gamma+z=c.\end{array}\right.$$

300. **K** = **R**

$$\left\{\begin{array}{l}x+y\cos\gamma+z\cos\beta=a\\ x\cos\gamma+y+z\cos\alpha=b\\ x\cos\beta+y\cos\alpha+z=c.\end{array}\right.$$

301. Resolver el sistema (**K** cualquiera)

$$\frac{x_1}{a_1}=\ldots=\frac{x_k}{a_k}=\ldots=\frac{x_n}{a_n},\qquad x_1+\ldots+x_n=s.$$

302. Resolver el sistema (**K** cuerpo conmutativo de característica p, se discutirá según el valor de p)

$$x_0 + x_1 = a_1, \ \ldots, \ x_0 + x_k = a_k, \ \ldots, \ x_0 + x_n = a_n, \ x_1 + \ldots + x_n = s.$$

303. *Estudiar el sistema de n ecuaciones con n incógnitas* (**K** = **C**), siendo la k-ésima ecuación la

$$x_1 + \ldots + x_{k-1} + \lambda x_k + x_{k+1} + \ldots + x_n = a_k.$$

a) *Utilizar la incógnita auxiliar* $s = x_1 + \ldots + x_n$.

b) Utilizar los determinantes.

c) Estudiar el sistema en que $(k = 1, \ldots, n)$, $a_k = (\lambda)^k$.

304. Estudiar el sistema de n ecuaciones con n incógnitas (**K** = **C**) en el que la k-ésima ecuación es

$$(a_1)^{k-1}x_1 + \ldots + (a_k)^{k-1}x_k + \ldots + (a_n)^{k-1}x_n = (a_{n+1})^{k-1}.$$

305. Estudiar el sistema de n ecuaciones con n incógnitas (**K** = **C**)

$$\begin{cases} (k = 1, 2, \ldots, n-1) \qquad x_k + x_{k+1} = 2a_k \\ \qquad\qquad\qquad\qquad\qquad x_n + x_1 = 2a_n. \end{cases}$$

Se estudiará primero el caso $n = 2$, $n = 3$, $n = 4$.

Aplicación: Determinar un polígono plano $M_1, M_2, \ldots, M_n, M_1$ conociendo los puntos medios de los lados sucesivos $A_1, A_2, \ldots, A_n$.

En los ejercicios 306 *al* 307 *se determinará una base del subespacio de* $\mathbf{K}^n$ *definido por el sistema considerado.*

306. $n = 3$

$$\begin{cases} x + 2y - z = 0 \\ 2x + 7y - 2z = 0 \\ -x + 3y + z = 0. \end{cases}$$

307. $n = 5$

$$\begin{cases} x + y - z - t + u = 0 \\ 2x + y - 4t + 4u = 0 \\ x + 2y - 3z + t - u = 0. \end{cases}$$

308. Si **K** es un cuerpo conmutativo de característica nula se considera el sistema $(m \leq n)$

$$(i = 1, 2, \ldots, m) \qquad \sum_{j=1}^{n} a_i^j x_j = 0 \qquad \text{con} \qquad a_i^j = i + j - 1$$

determinar una base del subespacio de $\mathbf{K}^n$ definido por este sistema.

En los ejercicios siguientes (309 al 312) *resultará fácil introducir el valor común de las razones como incógnita auxiliar* (se tomará K = C).

309. $$\frac{x + \lambda y + \mu z}{x} = \frac{y + \lambda z + \mu x}{y} = \frac{z + \lambda x + \mu y}{z}.$$

310. $$\frac{y + z - x}{a} = \frac{z + x - y}{b} = \frac{x + y - z}{c}.$$

311. $$\frac{bx + cy + h}{a} = \frac{cx + ay + h}{b} = \frac{ax + by + h}{c}.$$

312. $\dfrac{l - bz + cy}{a} = \dfrac{m - cx + az}{b} = \dfrac{n - ay + bx}{c}$.

313. Se considera el sistema ($\mathbf{K} = \mathbf{C}$)

$$\begin{cases} & ax_1 - x_2 = 0 \\ (k = 2, 3, \dots, n-1) & -x_{k-1} + ax_k - x_{k+1} = 0 \\ & -x_{n-1} + ax_n = 0 \end{cases}$$

se designa por D_n el determinante del sistema:

a) Demostrar que D_n se expresa en función de D_{n-1}, D_{n-2};

b) Demostrar que D_n es solución de la ecuación de recurrencia

$$u_n = au_{n-1} - u_{n-2} \qquad \text{con} \qquad u_0 = 1,\ u_1 = a.$$

Calcular D_n en función de n y a (utilizar el ej. 149, cap. 7);

c) Encontrar los valores de a que anulan D_n. Para cada uno de estos valores calcular $x_2, x_3, \dots, x_n$ en función de x_1.

Comparar los valores de x_p y x_q cuando $p + q = n + 1$.

314. Se considerán los dos sistemas

$$(S) \begin{cases} bz - cy = l \\ cx - az = m \\ ay - bx = n \end{cases} \qquad (S') \begin{cases} b'x - c'y = l' \\ c'x - a'y = m' \\ ay' - bx' = n'. \end{cases}$$

a) ¿En qué condición (S) tiene soluciones? Satisfecha esta condición y siendo x_0, y_0, z_0 una solución demostrar que todas las soluciones de (S) son de la forma (λ arbitrario)

$$x = x_0 + a\lambda, \qquad y = y_0 + b\lambda, \qquad z = z_0 + c\lambda.$$

b) ¿En qué condición (S) y (S′) tienen una solución común?

315. Se supone $\mathbf{K} = \mathbf{R}$ y $a > 0$, $b > 0$, $c > 0$, $d > 0$, demostrar que el sistema siguiente es imposible

$$\begin{cases} x + y + z + t = a \\ x - y - z + t = b \\ -x - y + z + t = c \\ -3x + y - 3z + 7t = d. \end{cases}$$

316. Se consideran m formas lineales $l^1, \dots, l^m$ definidas sobre $\mathbf{R}^n$:

a) Demostrar que, si hay n números $\lambda_1, \dots, \lambda_m$ *estrictamente positivos* tales que

$$\lambda_1 l^1 + \lambda_2 l^2 + \dots + \lambda_m l^m = 0$$

el sistema de las m inecuaciones

$$(S) \qquad (k = 1, 2, \dots, m) \qquad l^k(x) > 0$$

es imposible.

b) *Si el rango de las formas es superior o igual a* $m - 1$ demostrar que la condición enunciada en la pregunta a) es igualmente necesaria para que el sistema (S) sea imposible.

(Se empezará por estudiar el ejercicio 315).

11

POLINOMIOS

I. Definiciones generales.
II. Estudio de **K**[X], **K** cuerpo conmutativo.
III. Estudio de **K**[X_1, ..., X_m], **K** cuerpo conmutativo.

Vamos a estudiar en este capítulo objetos llamados *polinomios,* veremos que las definiciones y las propiedades de las operaciones algebraicas definidas sobre ellos son muy análogas a las de los «polinomios» estudiados en Matemáticas elementales. Veremos la causa a lo largo de nuestro estudio.

I. Definiciones generales

182. Polinomios con una indeterminada

a) Sea A un anillo unitario y conmutativo cuyos elementos neutros se representarán, cuando no haya lugar a confusión, por 0 y 1. (En la mayoría de los casos empleados A será el cuerpo **R** o el cuerpo **C**.)

Definición 1. — *Se llama* polinomio de una indeterminada[35], *con coeficientes en el anillo conmutativo unitario* A, *toda sucesión f de elementos a_i de* K, *es decir,* $(a_0, \dots, a_i, \dots)$ *todos nulos a partir de un cierto rango. Los elementos a_i de* K *son los* coeficientes *del polinomio f.*

Un polinomio (de una indeterminada) es, pues, una aplicación de **N** en A en que sólo un número finito de valores son no nulos; de ello resulta la definición de igualdad de dos polinomios

$$(a_0, \dots, a_i, \dots) = (b_0, \dots, b_i, \dots) \Rightarrow (\forall i \in \mathbf{N}) a_i = b_i.$$

b) En el conjunto de los polinomios de una indeterminada y de coe-cientes en A, que designaremos provisionalmente por $\mathcal{F}$, definiremos dos op-raciones internas, *suma* y *multiplicación* mediante las fórmulas siguientes:

$$f = (a_0, \dots, a_i, \dots), \qquad g = (b_0, \dots, b_j, \dots)$$

(35) El origen de esta denominación se dará en el párrafo siguiente.

(1) $$f + g = (a_0 + b_0, \ldots, a_i + b_i, \ldots)$$
(2) $$fg = (c_0, \ldots, c_k, \ldots)$$
con
$$c_k = a_0 b_k + a_1 b_{k-1} + \ldots + a_k b_0 = \sum_{i+j=k} a_i b_j.$$

La *adición* (1) proporciona a $\mathfrak{F}$ una *estructura de grupo abeliano*; es, por otra parte, la estructura del grupo $\mathfrak{F}(\mathrm{N}, \mathrm{A})$, si A se le considera como un grupo aditivo.

El opuesto de $f = (a_0, \ldots, a_i, \ldots)$ es $-f = (-a_0, \ldots, -a_i, \ldots)$, existe un *polinomio cero*: $0 = (0, \ldots, 0, \ldots)$.

Siendo A un anillo conmutativo, la *multiplicación* (2) es *conmutativa*, siendo A unitario existe un *polinomio unidad*: $e = (1, 0, \ldots, 0, \ldots)$; se tiene, en efecto, para todo polinomio f
$$fe = ef = f.$$

Un cálculo fácil demuestra que la *multiplicación* es *distributiva* con relación a la suma. Veamos que la multiplicación es *asociativa* sea
$$f = (a_0, \ldots, a_i, \ldots), \qquad g = (b_0, \ldots, b_i, \ldots), \qquad h = (c_0, \ldots, c_i, \ldots),$$
pongamos
$$p = fg = (p_0, \ldots, p_i, \ldots), \qquad q = gh = (q_0, \ldots, q_i, \ldots)$$
$$r = (fg)h = ph = (r_0, \ldots, r_i, \ldots), \qquad s = f(gh) = fq = (s_0, \ldots, s_i, \ldots)$$
tendremos
$$r_n = \sum_{h+k=n} p_h c_k = \sum_{h+k=n} \left(\sum_{i+j=h} a_i b_j \right) c_k$$
$$= \sum_{h+k=n} \left(\sum_{i+j=h} a_i b_j c_k \right) = \sum_{i+j+k=n} a_i b_j c_k$$
para pasar de la segunda suma a la tercera utilizamos la distributividad en A y para pasar de la tercera a la cuarta utilizamos también la distributividad en A y observamos que todo término de la tercera suma pertenece a la cuarta y recíprocamente; del mismo modo
$$s_n = \sum_{i+l=n} a_i q_l = \sum_{i+l=n} a_i \left(\sum_{j+k=l} b_j c_k \right)$$
$$= \sum_{i+l=n} \left(\sum_{j+k=l} a_i b_j c_k \right) = \sum_{i+j+k=n} a_i b_j c_k$$
luego para todo n, $r_n = s_n$ y cualesquiera que sean los polinomios f, g, h
$$(fg)h = f(gh).$$

El conjunto $\mathfrak{F}$ provisto de las operaciones (1) *y* (2) *es, pues, un anillo conmutativo unitario.*

Sea $\mathcal{F}_0$ el subconjunto de $\mathcal{F}$ descrito por $(a, 0, \ldots, 0, \ldots)$, cuando a describe A; la aplicación φ de A en $\mathcal{F}_0$ definida por

$$a \to \varphi(a) = (a, 0, \ldots, 0, \ldots)$$

es visiblemente *biyectiva*. Por otra parte $(a, b \in A)$,

$$\varphi(a+b) = (a+b, 0, \ldots, 0, \ldots) = (a, 0, \ldots, 0, \ldots) + (b, 0, \ldots) = \varphi(a) + \varphi(b)$$
$$\varphi(ab) = (ab, 0, \ldots, 0, \ldots) = (a, 0, \ldots, 0, \ldots)(b, 0, \ldots, 0, \ldots) = \varphi(a)\varphi(b)$$

resulta que $\mathcal{F}_0$ es un subanillo de A isomorfo al anillo A. Concluimos:

TEOREMA 1. — *Si* A *es un anillo conmutativo y unitario el conjunto* $\mathcal{F}$ *de los polinomios de una indeterminada y con coeficientes en* A, *provisto de la adición* (1) *y de la multiplicación* (2) *tiene una estructura de anillo conmutativo unitario, en el que el subanillo descrito por* $(a, 0, \ldots, 0, \ldots)$ *(a describiendo* A) *es isomorfo a* A.

c) Identificaremos el elemento a del anillo A y el polinomio $(a, 0, \ldots)$. Resulta de ello, primero, que el polinomio cero y el polinomio unidad se representarán, respectivamente, por 0 y 1. Como consecuencia de esta identificación A es un *subanillo* de $\mathcal{F}$. Por otra parte, podemos así definir sobre $\mathcal{F}$ una multiplicación externa, en la que el dominio de operadores es el anillo A, poniendo para todo λ de A y todo f de $\mathcal{F}$

$$\lambda f = (\lambda, 0, \ldots)(a_0, \ldots, a_i, \ldots) = (\lambda a_0, \ldots, \lambda a_i, \ldots). \tag{3}$$

Se verifica inmediatamente que en $\mathcal{F}$ provisto de la adición (1) y de esta multiplicación externa, se tiene cualesquiera que sean f y g de $\mathcal{F}$ y λ y μ de A

$$\begin{cases} \lambda(\mu f) = (\lambda\mu)f, & 1f = f \\ (\lambda+\mu)f = \lambda f + \mu f, & \lambda(f+g) = \lambda f + \lambda g. \end{cases} \tag{4}$$

OBSERVACIONES

1. Estas últimas igualdades, unidas al hecho de que $\mathcal{F}$ es un grupo abeliano para la adición demuestran que $\mathcal{F}$ provisto de la adición y de la multiplicación externa posee una estructura de A-módulo (ver ej. 179, final del capítulo 7).

2. Se observará que si la adición de los polinomios y la multiplicación externa son las operaciones habituales definidas sobre $\mathfrak{F}(E, A)$ (E conjunto cualquiera, A anillo conmutativo) no ocurre lo mismo con la multiplicación; la multiplicación de dos sucesiones (a_n), (b_n) (elementos de $\mathfrak{F}(N, A)$) está definida por $(a_n)(b_n) = (a_n b_n)$ (ver § 90, ej. 5); fórmula muy diferente de la fórmula (2), el producto de dos polinomios que hemos definido es un caso particular de lo que se llama en estudios ulteriores un producto de composición. Luego si el conjunto de los polinomios no provisto de operaciones es idéntico al conjunto de las sucesiones (a_n) de elementos de A (siendo a_i nulo a partir de cierto rango), no sucede lo mismo cuando se ha definido la multiplicación de los polinomios.

183. Noción de indeterminada. Notación A[X]

Consideremos el polinomio siguiente (δ_{ki} símbolo de KRONECKER)

$$e_k = (\delta_{k0}, \ldots, \delta_{ki}, \ldots) = (0, \ldots, 0, 1, 0, \ldots)$$

el único coeficiente no nulo de e_k, el 1, está en el $(k+1)$-ésimo lugar. Para todo polinomio f de $\mathcal{P}$ tendremos

$$f = (a_0, \ldots, a_k, \ldots) = \sum_k a_k e_k \tag{5}$$

donde la Σ se extiende a todos los índices k tales que $a_k \neq 0$; tenemos, pues, una suma de un número finito de términos. Además, observando que

$$f = \sum_k a_k e_k = 0 \Leftrightarrow (\forall\, k \in \mathbf{N}) \qquad a_k = 0$$

tendremos, si $f = (a_0, \ldots, a_k, \ldots)$, $g = (b_0, \ldots, b_k, \ldots)$,

$$f = g \Leftrightarrow f - g = 0 \Leftrightarrow (\forall\, k \in \mathbf{N}) \qquad a_k = b_k$$

luego la representación de f dada por el segundo miembro de (5) es única. Calculemos

$$e_p e_q = \sum_k a_k e_k$$

$$a_k = \sum_{i+j=k} \delta_{pi} \delta_{qj}$$

el único caso en que $\delta_{pi}\delta_{qj} \neq 0$ es aquel en que $(i, j) = (p, q)$, de donde

$$e_p e_q = e_{p+q}.$$

Designemos por X el polinomio e_1; tenemos, para todo entero natural $k \neq 0$: $e_k = \mathrm{X}^k$ (k es un exponente); por otra parte (§ 182, c), $e_0 = 1$, de donde:

TEOREMA 2. — *Designando con* X *el polinomio* (0, 1, 0, ..., 0, ...):

1. *Cualquiera que sea el entero* n *la relación* ($a_k \in \mathrm{A}$)

$$a_0 + a_1\mathrm{X} + \ldots + a_k\mathrm{X}^k + \ldots + a_n\mathrm{X}^n = 0$$

implica que $a_k = 0$ *para todo* K.

2. *Todo polinomio* f *de* $\mathcal{P}$ *se escribe de una manera única*

$$f = a_0 + \ldots + a_k\mathrm{X}^k + \ldots + a_n\mathrm{X}^n,$$

donde $a_0, \ldots, a_n$ *son los elementos de* A *verificando* $a_n \neq 0$.

El anillo $\mathcal{P}$ está engendrado por su parte $\mathrm{A} \cup \{\mathrm{X}\}$, lo representaremos A[X]; hubiéramos podido representar el polinomio e_1 por otra letra no empleada en este estudio, por ejemplo, Y; se ve inmediatamente que los anillos A[X] y A[Y] son isomorfos. Diremos que X es una *indeterminada,* lo que justifica *a posteriori* la denominación de *polinomio de una indeterminada* dada desde el principio de este estudio a los elementos de $\mathcal{P}$. *Diremos igualmente que todo elemento de* A[X] *es un polinomio en* X *con coeficientes en* A.

Cuando se escribe un elemento f de A[X] en la forma $f = a_0 + \ldots + a_nX^n$ (resp. $f = a_nX^n + \ldots + a_0$) se dice que f *está ordenado según las potencias crecientes* (resp. *decrecientes*) (*se sobrentiende de la indeterminada* X).

OBSERVACION

De una manera más general sea A un anillo conmutativo y α un elemento que no pertenece a A, de un superanillo B de A, no necesariamente conmutativo, pero que α permute con todo elemento de A.

Designemos por A[α] el subanillo de B engendrado por A $\cup$ {α}; A[α] estará descrito por los elementos de la forma $\Sigma a_k\alpha^k$, donde los elementos a_k de A son nulos, salvo un número finito de ellos. Pueden darse dos casos:

1.° Existe al menos una relación ($k = 1, \ldots, p$, $a_k \in$ A)

$$a_0 + a_1\alpha + \ldots + a_p\alpha^p = 0.$$

No siendo nulos los a_k, se dice entonces que α es *algebraico sobre el anillo* A.

2.° Cualquiera que sea n ($a_k \in$ A), la igualdad

$$a_0 + a_1\alpha + \ldots + a_n\alpha^n = 0$$

implica $a_k = 0$ *para todo entero* k, *se dice entonces que* α *es trascendente sobre* A. Vemos, pues, que la indeterminada X (polinomio e_1) de A[X] es trascendente sobre A. (A[α] y A[X] son, por otra parte, dos anillos isomorfos).

EJEMPLOS Y EJERCICIOS

1. $\sqrt{2}$ es algebraico sobre **Q** y **Z**; i es algebraico sobre **R**, **Q** y **Z**.

2. Si K es un cuerpo conmutativo y si α es un elemento de un supercuerpo L de K, si α no pertenece a K y permuta con todo elemento de K, demostrar que las dos propiedades siguientes son equivalentes:

a) α es algebraico sobre K.

b) K[α] es un subespacio vectorial de dimensión finita sobre K.

Si n es la dimensión del subespacio vectorial K[α] sobre K se dice que α es *algebraico de grado n sobre* K y que K[α] es una extensión algebraica de grado n sobre K, **Q**[$\sqrt{2}$] y **R**[i] (ver ej. 96 y 97, capítulo 5), son dos extensiones cuadráticas (es decir, de grado 2), respectivamente, de **Q** y de **R**.

3. Se llama *número algebraico de grado n*, todo número complejo algebraico de grado n sobre **Q** (o sobre **Z**, que es lo mismo) (ver ej. 151, capítulo 7). Se llama *número trascendente* todo número complejo trascendente sobre **Q** (o **Z**); por ejemplo, se ha demostrado en el último tercio del siglo XIX que e (base de los logaritmos neperianos) y π son trascendentes.

¿Cuál es el grado de $j = -\dfrac{1}{2} + i\dfrac{\sqrt{3}}{2}$ sobre **Q**, de $x = \sqrt[3]{2} + \sqrt[2]{3}$ sobre **Q**?

184. Grado de un polinomio con una indeterminada

a) Noción de grado

DEFINICIÓN 2. — *Dado el polinomio* $f = \Sigma a_hX^h$ de A[X], *siendo* A *un anillo conmutativo unitario, se llama* grado *de* f, *lo que se representa* grd f, *el mayor entero* n *tal que* $a_n \neq 0$.

Se deduce de esta definición que el *polinomio cero no tiene grado.*

A los polinomios $a_h X^h$ $(a_h \neq 0)$ se les llama *monomio* de grado h.

Los polinomios de grado 0 se llaman los *polinomios constantes* (veremos la causa en el § 185, *a*): por la identificación efectuada en el § 182, *c*) éstos son los elementos no nulos de A.

Si grd $f = n$, el monomio $a_n X^n$ se llama *monomio dominante* de f y a_n el *coeficiente dominante* de f; si $a_n = 1$ se dice que f es un *polinomio unitario.*

Sea f y g dos polinomios no nulos, pongamos

$$f = a_0 + \ldots + a_n X^n \qquad (a_n \neq 0)$$
$$g = b_0 + \ldots + b_p X^p \qquad (b_p \neq 0)$$

si $n \neq p$: resulta que $f + g \neq 0$ y grd $(f + g) = \sup (n, p)$; si $n = p$ y si $f + g \neq 0$, se tiene grd $(f + g) \leq \sup (n, p)$ (se tendría la igualdad si y sólo si $a_n + b_n \neq 0$).

Consideremos ahora el producto

$$fg = c_0 + \ldots + c_q X^q$$

la fórmula (2) del § 182 muestra que

$$q > n + p \Rightarrow c_q = 0, \qquad c_{n+p} = a_n b_p.$$

En el caso general, es decir, si el anillo A posee divisores de cero, no podemos afirmar que $c_{n+p} \neq 0$: puede suceder también que el polinomio fg sea nulo (ver más abajo ej. 1).

De donde resumiendo todos estos resultados:

TEOREMA 3.—*Sean f y g dos polinomios no nulos de* A[X] (A *anillo unitario conmutativo*):

1. *Si* grd $f \neq$ grd g *se tiene*

$$f + g \neq 0 \quad \text{y} \quad \text{grd } (f + g) = \sup (\text{grd } f, \text{ grd } g),$$

si grd $f =$ grd g *y si* $f + g \neq 0$ *se tiene*

$$\text{grd } (f + g) \leq \sup (\text{grd } f, \text{ grd } g).$$

2. *Si* $fg \neq 0$ *se tiene*

$$\text{grd } (fg) \leq \text{grd } f + \text{grd } g.$$

b) Caso particular en que A es un anillo íntegro

En este caso con las notaciones anteriores, $a_n \neq 0$ y $b_p \neq 0$ implica $c_{n+p} \neq 0$, luego $f \neq 0$ y $g \neq 0$ implica $fg \neq 0$, de donde:

TEOREMA 4.—*Si* A *es un anillo íntegro unitario, también lo es* A[X].
Además, si $f \neq 0$ *y* $g \neq 0$

$$\text{grd } (fg) = \text{grd } f + \text{grd } g.$$

Determinemos los polinomios inversibles de A[X] (A anillo íntegro unitario) se tendrá

$$fg = 1 \Rightarrow \text{grd } f + \text{grd } g = 0,$$

luego grd $f =$ grd $g = 0$, de donde:

COROLARIO. — *Si* A *es un anillo unitario, íntegro, los únicos elementos inversibles de* A[X] *son los elementos inversibles de* A.

OBSERVACIONES

1. Si se hubiera dado al polinomio cero un grado nulo, para todo f la relación $0f = 0$ daría $0 +$ grd $f = 0$.

2. El teorema 1 (§ 182) y el teorema 4 anterior muestran que si A es un anillo, respectivamente, conmutativo, unitario, íntegro, ocurre lo mismo con el anillo A[X]. Tales enunciados que «transfieren» ciertas propiedades del anillo A al anillo A[X] se llaman *teoremas de transferencia.*

EJERCICIOS

1. En A[X] ($A = \mathbf{Z}/6\mathbf{Z}$) dar ejemplos en que grd $(fg) <$ grd $f +$ grd g; dar ejemplos de divisores de cero.

2. Siendo A un anillo conmutativo unitario dotado de divisores de cero, ¿cuáles son los elementos inversibles del anillo A[X]?

185. Función polinómica de una variable. Sustitución de un polinomio en un polinomio

a) Función polinómica

Sea $f = a_0 + a_1X + \ldots + a_nX^n$ un elemento de A[X] (siendo A un anillo conmutativo unitario), podemos hacer corresponder a toda pareja (x, f) elemento de A × A[X] el elemento de A

$$a_0 + a_1x + \ldots + a_nx^n$$

que designaremos $f(x)$: diremos que $f(x)$ se ha obtenido *sustituyendo* por el elemento x de A el X en el polinomio f; definiremos así una aplicación de A en A:

DEFINICIÓN 3. — *A todo polinomio* $f = a_0 + a_1X + \ldots + a_nX^n$ *de* A[X] (A *anillo conmutativo unitario*) *se puede hacer corresponder una aplicación* $\tilde{f}$ *de* A *en* A *definida por*

$$(\forall x \in A) \qquad \tilde{f}(x) = a_0 + a_1x + \ldots + a_nx^n = f(x)$$

llamada función polinómica asociada *al polinomio* f.

El conjunto de las funciones polinómicas que aplican A en A, y que representaremos $\mathfrak{F}(A, A)$, es evidentemente un subanillo de $\mathfrak{F}(A, A)$ (ver § 90, ej. 5).

Cualesquiera que sean los polinomios f y g de A[X] y el elemento λ de A, las reglas de cálculo en el anillo A, muestran que para todo x de A

$$(f + g)(x) = f(x) + g(x), \qquad (fg)(x) = f(x)g(x), \qquad (\lambda f)(x) = \lambda f(x)$$

luego

$$\left(\widetilde{f + g}\right) = \tilde{f} + \tilde{g}, \qquad \widetilde{fg} = \tilde{f}\tilde{g}, \qquad \widetilde{\lambda f} = \lambda \tilde{f}.$$

Las dos primeras igualdades prueban que la aplicación φ de A[X] en $\mathfrak{F}$(A, A), definida por $\varphi(f) = \tilde{f}$ es un *homomorfismo* de anillos; al estar $\mathfrak{F}$(A, A) descrito por $\tilde{f}$ asociada a f, este homomorfismo es *suprayectivo, pero en general no es inyectivo* (ver ej. 1 más abajo). Indicaremos en el § 192 un caso en que este homomorfismo es un isomorfismo (lo que sucede con A = **R** como lo recuerda la observación siguiente).

A los polinomios identificados a los elementos de A están asociadas las *funciones constantes* de $\mathfrak{F}$(A, A), de donde que el nombre de *polinomios constantes* dado a los polinomios $\lambda (\lambda \in A)$ sea un abuso de lenguaje.

OBSERVACION

Supongamos A = **R**. En elemental se definían dos nociones distintas: *identidad* y *equivalencia* de dos polinomios y se demostraba (o se admitía) que cada una de ella implicaba la otra: se trata, en efecto, de dos igualdades, pero en dos conjuntos diferentes la *identidad* es la igualdad en **R**[X], la *equivalencia* es la igualdad en $\mathfrak{F}$(**R**, **R**). Se dice también que algunas veces que $f(X) = g(X)$ es una *igualdad formal* y $(\forall x \in \mathbf{R})$ $f(x) = g(x)$ una *igualdad numérica*; de donde el nombre de «polinomio formal» que dan ciertos autores a los polinomios: no lo haremos aquí, los términos «polinomio» y «función polinómica» son suficientemente claros para no tener que introducir un tercer término.

b) Prolongación de una función polinómica a un superanillo de A

Podemos prolongar $\tilde{f}$ a un superanillo conmutativo B y A, la prolongación (ver § 16) de $\tilde{f}$ a B tiene por valor $f(x)$ definida por (siendo $a_0, \dots, a_n$ elementos de A)

$$(\forall x \in B) \qquad f(x) = a_0 + a_1 x + \dots + a_n x^n \in B.$$

Diremos también que $f(x)$ se ha obtenido al sustituir por el elemento x de B la indeterminada X en f elemento de A[X].

En particular se puede tomar B = A[X] que es un superanillo conmutativo de A (§ 182, *c*), luego: $f = a_0 + a_1 X + \dots + a_n X^n \in A[X]$, $g \in A[X]$ implican

$$f(g) = a_0 + a_1 g + \dots + a_n g^n.$$

Se dice que se ha sustituido (a la indeterminada X*) por el polinomio g en el polinomio f.*

En particular se puede sustituir el polinomio $g = X$ en f, luego todo polinomio de A[X] puede escribirse

$$f = f(X).$$

OBSERVACION

Se puede igualmente suponer B no conmutativo a condición de no sustituir más que los elementos x *que permutan con todos los elementos de* A (ver ej. 5 más abajo).

EJEMPLOS Y EJERCICIOS

1. $A = \mathbf{Z}/2\mathbf{Z}$; al polinomio $f = X^2 - X$ está asociada una función polinómica definida sobre A, nula para todo elemento de A, y, sin embargo, $X^2 - X$ no es el polinomio nulo.

Formar un polinomio de coeficientes en $A = \mathbf{Z}/n\mathbf{Z}$ no nulo y tal que la función polinómica asociada sea nula para todo elemento de A.

2. Demostrar que en general f y g siendo dos polinomios de A[X], $f(g) \neq g(f)$.

3. Demostrar que grd $[f(g)] \leq$ (grd f) (grd g); si A es íntegro hay igualdad.

4. Si $\tilde{f}, \tilde{g}, \tilde{h}$ son las funciones polinómicas (aplicando A en A) asociadas, respectivamente, a los tres polinomios $f, g, h = f(g)$ de A[X], demostrar que $\tilde{h} = \tilde{f} \circ \tilde{g}$.

5. Sea K un cuerpo conmutativo, E un espacio vectorial sobre K y $B = \mathcal{L}(E)$ el anillo de los endomorfismos de E. Demostrar que si f es un endomorfismo de E, el subanillo B′ de B engendrado por K y f es conmutativo. Deducir que a todo polinomio $P = a_0 + a_1X + \dots + a_nX^n$ de K[X] se puede hacer corresponder el endomorfismo de E

$$P(f) = a_0 + a_1f + \dots + a_nf^n,$$

con para $n > 1$, $f^n = f^{n-1} \circ f$. Si P y Q son dos elementos de K[X] y PQ su producto, demostrar que

$$(PQ)(f) = P(f) \circ Q(f).$$

Rehacer el estudio precedente con $B = M_n(K)$, reemplazando f por $M(f)$.

186. Polinomios con varias indeterminadas

a) Anillo $A[X_1, \dots, X_p]$

Sea A un anillo conmutativo unitario, $A_1 = A[X]$ es igualmente un anillo conmutativo unitario; consideremos $A_2 = A_1[Y]$, es decir, el anillo de los polinomios con una indeterminada y con coeficientes en A_1. Hemos designado la nueva indeterminada por Y, se la podría representar por otra letra, salvo X, ya empleada (ver § 183, *a*). En A[X], por identificación, hemos obtenido $A \subset A[X]$, tendremos igualmente por identificación

$$A \subset A_1 = A[X] \subset A_2 = A_1[Y]$$

luego los elementos cero y unidad de A_2 son, respectivamente, iguales a los elementos cero y unidad de A_1, luego de A, es decir, *0* y 1. Todo elemento de A_2 puede escribirse

$$f = \sum_j f_j(X)Y^j = \sum_j\Big(\sum_i a_{ij}X^i\Big)Y^j = \sum_i\sum_j a_{ij}X^iY^j$$

donde los $f_j(X)$ (elementos de $A_1 = A[X]$) son nulos, salvo un número finito de ellos, lo mismo que los a_{ij} (elementos de A). El último miembro de las igualdades precedentes se ha obtenido aplicando las reglas de cálculo en un anillo (ver § 91, *e*). Siempre según estas reglas, siendo A_2 conmutativo

$$f = \sum_i \left(\sum_j a_{ij} Y^j \right) X^i = \sum_i g_i(Y) X^i$$

con $g_i(Y) = \sum_j a_{ij} Y^j$, perteneciendo a $A'_1 = A[Y]$; acabamos de probar que todo elemento de $A_2 = A_1[Y]$ es elemento de $A'_2 = A'[X]$ con $A_1 = A[X]$, $A'_1 = A[Y]$, se probará igualmente que todo elemento de A'_2 es elemento de A_2, luego $A_2 = A'_2$; representaremos estos anillos idénticos por $A[X, Y]$.

TEOREMA Y DEFINICIÓN 1'. — *Si* A *es un anillo conmutativo unitario, los dos anillos conmutativos unitarios* $A[X][Y]$ *y* $A[Y][X]$ *son iguales, se les representa por* $A[X, Y]$. *Este último anillo se llama anillo de los polinomios con dos indeterminadas y de coeficiente en* A.

$f = 0$ implica para todo j de **N**, $f_j(Y) = 0$ (o para todo i de **N**, $g_i(X) = 0$), es decir, para todo (i, j) de $\mathbf{N}^2$, $a_{ij} = 0$, de donde:

TEOREMA 2'. — *En el anillo* $A[X, Y]$, *siendo los* a_{ij} *elementos de* A *nulos, salvo un número finito, es*

1. $$\sum_i \sum_j a_{ij} X^i Y^j = 0 \Rightarrow (\forall\, i,\ j \in \mathbf{N}) a_{ij} = 0.$$

2. *Todo elemento f de* $A[X, Y]$ *se escribe de una manera única*

$$f = \sum \sum a_{ij} X^i Y^j.$$

OBSERVACION

Sea B un superanillo de A no forzosamente conmutativo y α y β dos elementos de B, no pertenecientes a A, pero permutando entre sí y permutando con todos los elementos de A, designemos por $A[\alpha, \beta]$ el subanillo de B engendrado por $A \cup \{\alpha, \beta\}$.

Si $\sum_i \sum_j a_{ij} \alpha^i \beta^j = 0$ ($a_{ij} \in A$) implica $a_{ij} = 0$ para toda pareja de enteros naturales (i, j) diremos que α y β son *algebraicamente independientes sobre* A. Por ejemplo, las indeterminadas X e Y son algebraicamente independientes sobre A (teorema 2'). Se ve fácilmente que si α y β son algebraicamente independientes sobre A, $A[\alpha, \beta]$ y $A[X, Y]$ son dos anillos isomorfos; $A[X, Y]$ es, pues, el anillo engendrado por $A \cup \{X, Y\}$.

Definiremos igualmente por recurrencia, el *anillo de los polinomios con p indeterminadas* $X_1, \ldots, X_p$ y con coeficientes en A poniendo

$$(k = 2, \ldots, p) \qquad A[X_1, \ldots, X_k] = A[X_1, \ldots, X_{k-1}][X_k]$$

se ve fácilmente que

$$A[X_1, \ldots, X_p] = A[X_{\sigma(1)}, \ldots, X_{\sigma(p)}] = A[X_{\sigma(1)}, \ldots, X_{\sigma(n)}][X_{\sigma(n+1)}, \ldots, X_{\sigma(p)}]$$

en donde σ es una permutación de $[1, p]$ y n tal que $1 \leq n < p$. Todo polinomio f de $A[X_1, \ldots, X_p]$ se escribe de una manera única

$$f = \sum_{i_1} \ldots \sum_{i_p} a_{i_1 \ldots i_p} X_1^{i_1} \ldots X_p^{i_p}$$

donde los coeficientes $a_{i_1 \ldots i_p}$ (elementos de A) son nulos, salvo un número finito de ellos y $f = 0$ implica $a_{i_1 \ldots i_p} = 0$ para toda p-étupla $(i_1 \ldots i_p)$.

b) Grados parciales. Grado total. Polinomios homogéneos

DEFINICIÓN 3. — *Dado* $f = \sum_{i_1} \ldots \sum_{i_p} a_{i_1 \ldots i_p} X_1^{i_1} \ldots X_p^{i_p}$ *se llama* grado parcial *de f, respecto a* X_m $(1 \leq m \leq p)$ *el mayor de los enteros* i_m, *para todas las p-étuplas* $(i_1, \ldots, i_p)$ *tales que* $a_{i_1 \ldots i_p} \neq 0$ *y* grado total *de f el mayor de los enteros* $i_1 + \ldots + i_p$ *para todas las p-étuplas* $(i_1, \ldots, i_p)$ *tales que* $a_{i_1 \ldots i_p} \neq 0$, *se representa* grd f *el grado total.*

El grado parcial de f respecto X_m es, en consecuencia, el grado de f considerado como polinomio en X_m y con coeficientes en $A[X_1, \ldots, X_{m-1}, X_{m+1}, \ldots, X_p]$ si $1 < m < n$, en $A[X_2, \ldots, X_p]$ si $m = 1$ y en $A[X_1, \ldots, X_{p-1}]$ si $m = p$.

Observemos que el polinomio 0 no tiene grado (ni total, ni parcial), las propiedades de los grados dadas en el teorema 3 del § 184 son también exactas para los grados totales y para los grados parciales respecto a una misma indeterminada.

Se dice que $a_{i_1 \ldots i_p} X_1^{i_1} \ldots X_p^{i_p}$ $(a_{i_1 \ldots i_p} \neq 0)$ es un *monomio,* su grado total es $i_1 + \ldots + i_p$. Todo polinomio es una suma de monomios.

DEFINICIÓN 4. — *Se dice que un polinomio de* $A[X_1, \ldots, X_p]$ *es* homogéneo *de grado total n si es suma de monomios de grado n.*

Sea f un polinomio de grado total n, designemos por h_m $(0 \leq m \leq n)$ la suma de sus monomios de grado m (h_m puede ser nulo, pero $h_n \neq 0$, si no grd $f < n$); se tiene, pues, y esto de una manera única,

$$f = h_0 + \ldots + h_m + \ldots + h_n$$

h_m se llama la *parte homogénea de grado m* del polinomio f.

Se puede ordenar un polinomio de $A[X_1, \ldots, X_p]$ de múltiples maneras. Hemos ya visto que todo elemento de $A[X, Y]$ puede ordenarse según las potencias crecientes o decrecientes de X (resp. Y), por ejemplo, para $p(X, Y)$ de segundo grado (a, b, c, d, e, f elementos de A)

$$p(X, Y) = aX^2 + (bY + c)X + dY^2 + cY + f$$
$$p(X, Y) = dY^2 + (bX + e)Y + aX^2 + cX + f.$$

Se puede también escribir

$$p(X, Y) = (aX^2 + bXY + dY^2) + (cX + eY) + f$$

y en general por f de $A[X_1, ..., X_p]$

$$f(X_1, ..., X_p) = \sum_k h_k(X_1, ..., X_p)$$

con h_k homogéneo de grado k en $X_1, ..., X_p$. Nos basta entonces ordenar los términos de un polinomio homogéneo.

Diremos que el monomio $\alpha X_1^{i_1} \dots X_p^{i_p}$ $(i_1 + ... + i_p = k)$ es *menos alto* (o de *altura más débil*) que el monomio $\beta X_1^{j_1} \dots X_p^{j_p}$ $(j_1 + ... + j_p = k)$ si $(i_1, ..., i_p)$ es anterior a $(j_1, ..., j_p)$ en el *orden lexicográfico* (ver § 24, *c*); siendo las indeterminadas las letras de un alfabeto $X_1, X_2, ..., X_p$ dadas en este orden los monomios

$$\alpha X_1^{i_1} X_2^{i_2} \dots X_p^{i_p} = \alpha \underbrace{X_1 X_1 \dots X_1}_{i_1 \text{ letras}} \underbrace{X_2 X_2 \dots X_2}_{i_2 \text{ letras}} \dots \underbrace{X_p \dots X_p}_{i_p \text{ letras}}$$

están también clasificados según su altura, como las palabras de un diccionario (naturalmente algunos de los enteros $i_1, ..., i_p$ pueden ser nulos).

Se constata inmediatamente que todo polinomio no nulo tiene un *monomio único de altura máxima y que si* A *es íntegro el monomio más alto del producto de dos polinomios homogéneos no nulos es el producto de los monomios más altos de cada uno de los dos factores.*

EJEMPLOS Y EJERCICIOS

1. Ordenar por orden de altura decreciente los polinomios homogéneos de grado 2, 3 o 4 de dos o tres indeterminadas.

2. Si X_{ij} $(1 \leq i \leq n, 1 \leq j \leq n)$ son n^2 indeterminadas se pone

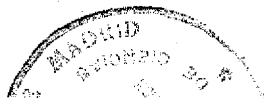
$$\Delta(X_{11}, ..., X_{nn}) = \sum_{p \in \mathfrak{S}_n} \varepsilon(p) X_{1p(1)} \dots X_{np(n)} = \begin{vmatrix} X_{11} & \dots & X_{n1} \\ \vdots & & \vdots \\ X_{1n} & \dots & X_{nn} \end{vmatrix}$$

$\Delta(X_{11}, ..., X_{nn})$ es un polinomio con n^2 indeterminadas con coeficientes en todo anillo conmutativo unitario A, se le llama el *determinante de las* n^2 *indeterminadas* X_{ij}, es homogéneo de grado n y es de grado parcial 1 respecto a cada indeterminada.

Se puede sustituir a k de estas indeterminadas X_{ij}, por k valores de A se obtiene un polinomio con $n^2 - k$ indeterminadas y con coeficientes en A.

3. ¿Cuál es el número de términos de $h_n(X, Y)$? (h_n homogéneo de grado n). Deducir el número de términos del polinomio $f(X, Y)$ de grado total n.

Las mismas preguntas para $h_n(X_1, ..., X_p)$, homogéneo de grado n y $f(X_1, ..., X_p)$ de grado n (utilizar el ejercicio 37, final del capítulo 2).



c) Caso en que el anillo A es íntegro

El teorema 4 del § 184 y un razonamiento por inducción nos permite enunciar:

TEOREMA 4. — *Si* A *es un anillo íntegro unitario, también lo son todos los anillos* $A[X_1, ..., X_p]$ *cualquiera que sea el entero* p.

Además, si f *y* g *son elementos no nulos de* $A[X_1, ..., X_p]$

$$\text{grd }(fg) = \text{grd }(f) + \text{grd }(g).$$

d) Función polinómica de varias variables

Sea $f = \sum_{i_1} \cdots \sum_{i_p} a_{i_1 \dots i_p} X^{i_1} \dots X^{i_p}$ un polinomio de $A[X_1, ..., X_p]$ (A anillo conmutativo unitario) y $(x_1, ..., x_p)$ un elemento de A^p, pondremos

$$f(x_1, ..., x_p) = \sum_{i_1} \cdots \sum_{i_p} a_{i_1 \dots i_p} x_1^{i_1} \dots x_p^{i_p}$$

diremos que hemos sustituido por x_i de A la indeterminada X_i para todo i de $[1, p]$. Definimos así una aplicación $\tilde{f}$ de A^p en A poniendo

$$(\forall (x_1, ..., x_p) \in A^p) \qquad \tilde{f}(x_1, ..., x_p) = f(x_1, ..., x_p) \in A$$

$\tilde{f}$ es una función de p variables llamada *función polinómica de* p *variables, asociada al polinomio* f. Designemos su conjunto por $\mathcal{F}(A^p, A)$; como en el caso de $p = 1$ se ve que φ, aplicación de $A[X_1, ..., X_p]$ en $\mathcal{F}(A^p, A)$ definida por $\varphi(f) = \tilde{f}$ es un *homomorfismo suprayectivo de anillos*.

Si $x_1, ..., x_p$ pertenecen a un superanillo conmutativo B del anillo A, podemos definir el elemento $f(x_1, ..., x_p)$ de B poniendo también

$$f(x_1, ..., x_p) = \sum_{i_1} \cdots \sum_{i_p} a_{i_1 \dots i_p} x_1^{i_1} \dots x_p^{i_p}$$

se dice también que $x_1, ..., x_p$ han sustituido, respectivamente, las indeterminadas $X_1, ..., X_p$ en el elemento f de $A[X_1, ..., X_p]$. Podemos en particular tomar $B = A[X_1, ..., X_p]$, luego dados $p + 1$ polinomios (elementos de B) $f, g_1, ..., g_p$ podemos definir $f(g_1, ..., g_p)$. Se puede tomar $g_h = X_h$ para $h = 1, 2, ..., p$, de modo que todo elemento f de $A[X_1, ..., X_p]$ puede escribirse

$$f = f(X_1, ..., X_p).$$

e) Propiedades de los polinomios homogéneos

Sea h_m un polinomio de grado m, elemento de $A[X_1, ..., X_p]$, siendo Y una $p+1$-ésima indeterminada, en $A[X_1, ..., X_p, Y]$ tendremos

$$h_m(X_1 Y, ..., X_p Y) = Y^m h_m(X_1, ..., X_p)$$

recíprocamente sea f un polinomio de $A[X_1, \ldots, X_p]$ tal que en el anillo $A[X_1, \ldots, X_p, Y]$ se tenga

$$f(X_1Y, \ldots, X_pY) = Y^n f(X_1, \ldots, X_p).$$

Si $f \neq 0$ el grado total de $f(X_1, \ldots, X_p)$ es igual al grado parcial de $f(X_1Y, \ldots, X_pY)$ relativamente a Y, luego a n. Escribamos (h_k homogéneo de grado k)

$$f = h_0 + \ldots + h_k + \ldots + h_n$$

obtendremos, teniendo en cuenta la igualdad dada y de la propiedad de h_k ($k = 0, \ldots, n$)

$$f(X_1Y, \ldots, X_pY) = \sum_{k=0}^{n} Y^k h_k(X_1, \ldots, X_p) = \sum_{k=0}^{n} Y^n h_k(X_1, \ldots, X_p),$$

luego $h_k = 0$ para todo k tal que $k \neq n$, de donde: *Un polinomio f, elemento de $A[X_1, \ldots, X_p]$ es homogéneo de grado n si y sólo si*

$$f(X_1Y, \ldots, X_pY) = Y^n f(X_1, \ldots, X_p)$$

en el anillo $A[X_1, \ldots, X_p, Y]$, *siendo* Y *una* $p + 1$-ésima *indeterminada.*

A todo polinomio f de grado n de $A[X_1, \ldots, X_p]$ asociemos el polinomio $F = \varphi(f)$ definido por (h_m es la parte homogénea de grado m de f)

$$F(X_1, \ldots, X_p, T) = \sum_{m=0}^{n} h_m(X_1, \ldots, X_p)\, T^{n-m}$$

F es un polinomio homogéneo, de grado n, *de* $A[X_1, \ldots, X_p, T]$, donde T es una $p + 1$-ésima variable llamada *variable de homogeneidad*. Si designamos por H_{p+1} el conjunto de los polinomios homogéneos con coeficientes en A y de $p + 1$ indeterminadas $X_1, \ldots, X_p$, T definimos así una aplicación

$$\varphi : A[X_1, \ldots, X_p] \to H_{p+1}.$$

Por otra parte, a todo polinomio $G(X_1, \ldots, X_p, T)$ de H_{p+1} asociemos el polinomio $g = \psi(G)$ de $A[X_1, \ldots, X_p, T]$ definido por

$$g(X_1, \ldots, X_p) = G(X_1, \ldots, X_p, 1)$$

definiremos así una aplicación

$$\psi : H_{p+1} \to A[X_1, \ldots, X_p].$$

Estas dos aplicaciones permiten, pues, pasar de un polinomio no homogéneo con p indeterminadas a un polinomio homogéneo de $p + 1$ indeterminadas y recíprocamente.

Desgraciadamente φ y ψ no son biyecciones y no son recíprocas una de otra (ver ej. siguiente).

EJERCICIOS

4. Se pone $f = X^2 + Y^2 + 2X - 3Y + 5$, calcular $F = \varphi(f)$ y $\psi(F)$. Se pone $G = T(X^2 + Y^2) + 2XT^2 - 3YT^2 + 5T^3$, calcular $g = \psi(G)$ y $\varphi(g)$.

5. Demostrar que $\psi \circ \varphi$ es la identidad de $A[X_1, ..., X_p]$, deducir que φ es inyectiva y ψ suprayectiva (ver § 15, ej. 1 y 4, fin del capítulo 1).

Demostrar que φ no es suprayectiva y que ψ no es inyectiva.

Determinar $\varphi(A[X_1, ..., X_p]) = H'_{p+1}$, deducir una biyección φ' de $A[X_1, ..., X_p]$ sobre H'_{p+1} y su recíproca $\psi' = (\varphi')^{-1}$.

II. Estudio de K[X], K cuerpo conmutativo

Designaremos, respectivamente, por 0 y por 1 el cero y la unidad del cuerpo K. Algunos de los resultados demostrados son válidos en casos más generales, lo señalaremos eventualmente en la teoría o en los ejercicios.

187. Espacio vectorial y álgebra K[X]

a) Siendo K un cuerpo conmutativo, las propiedades traducidas por las relaciones (4) (§ 182) y el hecho de que, para la adición, K[X] es un grupo abeliano, muestran que K[X] tiene una estructura de espacio vectorial sobre K. Por otra parte, el teorema 2 (§ 183) muestra que $\{1, X, ..., X^n, ...\}$ es una base del espacio vectorial K[X]: es una familia libre y engendra K[X].

Finalmente cualesquiera que sean los polinomios f, g de K[X] y el elemento λ del cuerpo K

$$(\lambda f)g = f(\lambda g) = \lambda(fg)$$

luego:

TEOREMA 5. — *Siendo* K *un cuerpo conmutativo, el conjunto* K[X] *provisto de las operaciones* $(f, g) \to (f + g)$ *y* $(\lambda, f) \to \lambda f$ *tiene una estructura de espacio vectorial sobre* K, *y es de dimensión infinita numerable;* $\{1, X, ..., X^n, ...\}$ *es una base llamada base canónica de* K[X].

Provisto, además, de la multiplicación $(f, g) \to fg$, K[X] *tiene una estructura de álgebra sobre* K.

Por abuso de notación, representamos por K[X] el *grupo aditivo*, el *anillo íntegro* (§ 184, teorema 4), el *espacio vectorial sobre* K y el *álgebra* sobre K de los polinomios con una indeterminada X y con coeficientes en el cuerpo conmutativo K.

b) Siendo K un cuerpo conmutativo, consideremos el subconjunto $\mathcal{P}_n$ de K[X] descrito por el polinomio nulo y los polinomios de grado inferior o igual a n.

Por abuso de lenguaje diremos que $\mathcal{E}_n$ es el *conjunto de los polinomios de grado inferior o igual a n.*

Cualesquiera que sean f y g de $\mathcal{E}_n$ y λ de K

$$f + g \in \mathcal{E}_n, \qquad \lambda f \in \mathcal{E}_n.$$

$\mathcal{E}_n$ es, en consecuencia, un subespacio vectorial de K[X]; por otra parte, está descrito por los polinomios $a_0 + a_1 X + \ldots + a_n X^n$. Está, pues, engendrado por los $n + 1$ polinomios 1, X, ..., X^n que son independientes, luego:

Corolario. — *Siendo* K *un cuerpo conmutativo, el conjunto de los polinomios de* K[X] *de grado inferior o igual a* n *es un subespacio vectorial de dimensión* $n + 1$ *del espacio vectorial* K[X].

c) Consideremos una sucesión finita estrictamente creciente de enteros positivos

$$i_1 < i_2 \ldots < i_h < \ldots < i_p$$

y p polinomios $f_{i_1}, \ldots, f_{i_p}$ no nulos tales que

$$(h = 1, \ldots, p) \qquad \operatorname{grd}(f_{i_h}) = i_h$$

que son independientes; en efecto, sea

$$f = \lambda_{i_1} f_{i_1} + \ldots + \lambda_{i_h} f_{i_h} + \ldots + \lambda_{i_p} f_{i_p} = 0$$

el coeficiente dominante de f es único: es el de $\lambda_{i_p} f_{i_p}$, luego $\lambda_{i_p} = 0$, resulta que

$$\lambda_{i_1} f_{i_1} + \ldots + \lambda_{i_{p-1}} f_{i_{p-1}} = 0$$

se tendrá, pues, igualmente $\lambda_{i_{p-1}} = 0, \ldots, \lambda_{i_h} = 0, \ldots, \lambda_{i_1} = 0$, luego:

Teorema 6. — *Siendo* K *un cuerpo conmutativo, toda familia de polinomios* f_h *no nulos de* K[X] *tales que* $\operatorname{grd}(f_h) = h$, *siendo distintos dos a dos los grados de los polinomios, es libre.*

Corolario. — *El espacio vectorial* $\mathcal{E}_n$ *de los polinomios de grado inferior o igual a* n, *con coeficientes en el cuerpo conmutativo* K, *admite por base toda familia* $f_0, \ldots, f_h, \ldots, f_n$ *de polinomios tales que* $\operatorname{grd}(f_h) = h$, *para todo* h *de* $[0, n]$.

EJERCICIOS

1. Siendo K un cuerpo de características nula demostrar que hay una familia única $a_0, \ldots, a_n$ de elementos de K tales que para todo polinomio f de grado $\leq n$ de K[X]

$$f(X) = a_0 + a_1 X + a_2 X(X - 1) + \ldots + a_n X(X - 1) \ldots (X - n + 1).$$

Calcular $a_0, \ldots, a_n$. Generalizar reemplazando 0, 1, ..., $n - 1$ por n elementos dos a dos distintos del cuerpo K.

2. Si a y b son dos elementos distintos de K demostrar que todo polinomio de grado $2n$ de K[X] se escribe de una manera única

$$f(X) = \sum_{k=-n}^{n} a_k(X-a)^{n+k}(X-b)^{n-k}.$$

3. En K[X] se consideran los polinomios A_0, A_1 de primer grado y linealmente independientes y los polinomios B_0, B_1, B_2 de segundo grado linealmente independientes, demostrar que A_0, A_1, $(B_0)^2$, B_0B_1, B_0B_2 son independientes. ¿Qué subespacio de K[X] engendran?

Generalizar (se observará que si P_0, P_1, ..., P_n son de grado n y linealmente independientes engendran el mismo subespacio de K[X] que $\{1, X, ..., X_n\}$).

188. Derivación en K[X]. Fórmula de Taylor

a) Derivación en K[X]

DEFINICIÓN 5. — *Dado un polinomio f de* K[X], K *cuerpo conmutativo, se llama* polinomio derivado *de f, y se representa $Df = f'$, el coeficiente de* Y *en el polinomio* $f(X+Y)$ *perteneciente al anillo* K[X][Y] *de los polinomios en* Y *con coeficientes en* K[X].

El polinomio $f^{(k)}$ definido por $f^{(1)} = f'$, ..., $f^{(k)} = [f^{(k-1)}]'$ $(k > 1)$ se llama polinomio derivado k-ésima *de f.*

Se tiene, pues, $D^k f = f^{(k)}$, con $D^1 = D$ y $D^k = D^{k-1} \circ D$ $(k > 1)$. Sea f un elemento de K[X], tendremos

$$f(X+Y) = g_0(X) + g_1(X)Y + \dots + g_p(X)Y^p$$

sustituyendo Y por 0 obtenemos $f(X) = g_0(X)$ y por definición $g_1(X) = f'(X)$, $(h(X, Y) \in K[X, Y])$

$$f(X+Y) = f(X) + f'(X)Y + h(X, Y)Y^2$$

dicho de otro modo, $f(X+Y) - f(X) - Yf'(X)$ es divisible por Y^2 en el anillo K[X, Y], escribiremos

$$(1) \qquad f(X+Y) \equiv f(X) + Yf'(X) \qquad (\text{mod } Y^2).$$

Recíprocamente, toda relación en K[X, Y]

$$f(X+Y) \equiv f(X) + Yf_1(X) \qquad (\text{mod } Y^2)$$

implica $f_1 = f'$, luego la relación (1) determina el polinomio derivado f' de f.

OBSERVACION

Supongamos $K = \mathbf{R}$ y sea $\tilde{f}$ la función polinómica asociada a f, tendremos $(x, y \in \mathbf{R})$

$$\tilde{f}(x+y) - \tilde{f}(x) = y\tilde{f}'(x) + \tilde{h}(x, y)y^2$$

lo que implica

$$\lim_{y \to 0} \frac{\tilde{f}(x+y) - \tilde{f}(x)}{y} = \tilde{f}'(x),$$

luego si $K = \mathbf{R}$ *la derivada de* $\tilde{f}$ *asociada a* f *es la función asociada al polinomio derivado tal como acabamos de definirlo.*

Sea un segundo polinomio g de K[X] y λ un elemento cualquiera de K, tendremos

$$(2) \qquad g(X+Y) \equiv g(X) + Yg'(X) \qquad (\text{mod } Y^2)$$

de donde

$$f(X+Y) + g(X+Y) \equiv f(X) + g(X) + Y[f'(X) + g'(X)] \qquad (\text{mod } Y^2)$$
$$\lambda f(X+Y) \equiv \lambda f(X) + Y[\lambda f'(X)] \qquad (\text{mod } Y^2)$$
$$f(X+Y)g(X+Y) \equiv f(X)g(X) + Y[f'(X)g(X) + f(X)g'(X)] \qquad (\text{mod } Y^2)$$

luego:

TEOREMA 7. — *Siendo* f, g *dos polinomios cualesquiera de* K[X] *y* λ *un elemento cualquiera del cuerpo* K:

1. $(f+g)' = f' + g'$, $(\lambda f)' = \lambda f'$, $(fg)' = f'g + fg'$.
2. *La aplicación* D *definida por* $f \to D(f) = f'$ *es un endomorfismo del espacio vectorial* K[X]; *también lo es* D^k *definida por* $D^1 = D$ *y* $D^k = D^{k-1} \circ D$ *para* $k > 1$.

Las fórmulas (1) muestran que D es una derivación en el álgebra K[X] tal como la hemos definido en el ejercicio 168 del capítulo 7.

Aplicando la definición y los resultados anteriores vemos que

$$(a, a_0, \ldots, a_n \in K \quad y \quad k \in \mathbf{N})$$
$$(a)' = 0,\ (X)' = 1,\ (X^2)' = 2X,\ \ldots,\ (X^k)' = kX^{k-1}$$
$$(a_0 + a_1X + \ldots + a_nX^n)' = a_1 + 2a_2X + \ldots + na_nX^{n-1}.$$

ATENCIÓN: La relación $f' = 0$ no caracteriza los polinomios constantes (0 y los polinomios de grado nulo); en efecto, si K es de *característica* $p \neq 0$, $(X^p)' = pX^{p-1} = 0$ (ver ej. 1). Supondremos en todo el resto de este párrafo salvo mención contraria— que K es de característica nula.

b) Grado de los polinomios derivados sucesivos. Fórmula de Taylor

Sea k un entero estrictamente positivo, se ve fácilmente por recurrencia

$$1 \leq h \leq k \Rightarrow (X^k)^{(h)} = k(k-1)\ldots(k-h+1)X^{k-h}$$
$$h > k \Rightarrow (X^k)^{(h)} = 0$$

en particular $(X^k)^{(k)} = k!$

La aplicación $f \to D^h(f) = f^{(h)}$ siendo un endomorfismo del espacio vectorial K[X], si K es de *característica nula* y f un polinomio de *grado* n

$$1 \leq h \leq n \Rightarrow \text{grd } f^{(h)} = n - h$$
$$h > n \Rightarrow f^{(h)} = 0.$$

Sea E el espacio vectorial sobre K descrito por los polinomios de grado

inferior o igual a n: dim $E = n + 1$ y $\{1, X - a, \ldots, (X - a)^n\}$ $(a \in K)$, es una base de E (ver § 187, *c*); luego para todo $f \in E$

$$f(X) = \lambda_0 + \lambda_1(X - a) + \ldots + \lambda_k(X - a)^k + \ldots + \lambda_n(X - a)^n$$

calculemos la k-ésima derivada de los dos miembros tendremos, siendo g un elemento de K[X],

$$f^{(k)}(X) = k!\,\lambda_k + (X - a)g(X).$$

Sustituyamos por a la X, obtenemos

$$f^{(k)}(a) = k!\,\lambda_k$$

de donde:

TEOREMA 8. — *Siendo* K *un cuerpo de característica nula, y* f *un elemento de* K[X] *de grado* n, *si* $1 \leq h \leq n$, $f^{(h)}$ *es no nulo y de grado* $n - h$, *si* $h > n$, $f^{(h)}$ *es nulo.*

Para todo $a \in K$ *y todo polinomio* f *de grado* n *de* K[X]

$$f(X) = f(a) + (X - a)f'(a) + \ldots + \frac{(X - a)^k}{k!} f^{(k)}(a) + \ldots + \frac{(X - a)^n f^{(n)}(a)}{n!}.$$

(fórmula de Taylor para los polinomios).

EJERCICIOS

1. ¿Cuál es el núcleo de D, endomorfismo del espacio vectorial K[X]? (Discutir según la característica p del cuerpo K.) ¿Cuál es el núcleo del endomorfismo D^k $(k > 1)$?
2. Extender la derivación de los polinomios al anillo A[X] (A anillo conmutativo unitario) utilizando la relación (1) en el anillo A[X, Y].
3. Demostrar la fórmula de TAYLOR sin utilizar la estructura de espacio vectorial de K[X]. Observar que para todo a de K y h de N, $X^h = (X - a + a)^h$. Aplicar la fórmula de TAYLOR a A[X] (A anillo conmutativo, unitario, de característica nula).
4. Sea $f \in K[X]$, $g \in K[Y]$ se pone $h(Y) = f(g)$, demostrar que $h'(Y) = f'(g)g'(Y)$.

189. División euclídea de los polinomios

a) Caso en que f es divisible por $g \neq 0$

Siendo K un cuerpo conmutativo, K[X] es un anillo unitario íntegro que se aplica el teorema 2 del § 98: si el polinomio f es divisible por $g \neq 0$ f pertenece al ideal principal (g) y existe q único tal que $f = gq$, siendo nulo si y sólo si $f = 0$. Supongamos $f \neq 0$, $g \neq 0$ y $f = gq$, se tendrá

$$\text{grd } q = \text{grd } f - \text{grd } g \geq 0$$

busquemos q por el *método de los coeficientes indeterminados*, suponiendo f y g ordenados respecto a las potencias decrecientes de X

$$f = a_n X^n + \ldots + a_0, \qquad g = b_m X^m + \ldots + b_0$$

con $a_n b_m \neq 0$ y $n \geq m$ (siendo las c_k incógnitas) y

$$q = c_{n-m}X^{n-m} + \dots + c_0 \qquad (c_{n-m} \neq 0)$$

tendremos

$$(a_nX^n + \dots + a_0) = (b_mX^m + \dots + b_0)(c_{n-m}X^{n-m} + \dots + c_0)$$

de donde

$$a_n = b_m c_{n-m} \Rightarrow c_{n-m} = a_n(b_m)^{-1}.$$

Consideremos los polinomios

$$f_1 = f - \frac{a_n}{b_m} X^{n-m} g, \qquad q_1 = q - \frac{a^n}{b_m} X^{n-m}$$

tenemos

$$f_1 = gq_1$$

o bien $f_1 = 0$ y entonces $q_1 = 0$, o bien $f_1 \neq 0$ y entonces $q_1 \neq 0$ y

$$\text{grd } q_1 = \text{grd } f_1 - \text{grd } g \geq 0$$

podemos calcular, como más arriba, el coeficiente dominante de q_1 que es el segundo coeficiente no nulo de q (ordenado con relación a las potencias decrecientes). Haciendo esta operación un número finito de veces obtendremos todos los coeficientes de q.

b) **Existencia y unicidad del cociente y del resto en la división euclídea**

Supongamos siempre f y g distintos de cero, pero f *no divisible por* g: no existe q tal que $f - gq = 0$, consideremos el conjunto de los polinomios $f - gp$ en donde p recorre K[X], hay entre todos ellos los polinomios p tales que $f - gp \neq 0$, luego los polinomios q tales que

$$(\forall p \in K[X]) \qquad \text{grd } (f - gq) \leq \text{grd } (f - gp);$$

en efecto, el conjunto de los grados de $f - gp$ es una parte de **N**; tiene, en consecuencia, un elemento menor. Sea l este grado mínimo y q uno de los polinomios tales que grd $(f - gq) = l$, pongamos

$$\begin{aligned} f &= a_nX^n + \dots + a_0 \qquad (a_n \neq 0) \\ g &= b_mX^m + \dots + b_0 \qquad (b_m \neq 0) \\ f - gq &= c_lX^l + \dots + c_0 \qquad (c_l \neq 0) \end{aligned}$$

Veamos que $l < m$; supongamos $l \geq m$, el polinomio

$$p = q + \frac{c_l}{b_m} X^{l-m}$$

es tal que

$$f - gp = \left(c_{l-1} - \frac{c_l}{b_m} b_{m-1} \right) X^{l-1} + \dots + c_0$$

luego $f - gq$ no sería de grado mínimo.

Resulta inmediatamente que q es único. Supongamos, en efecto, que existe q y q_1 tales que

$$q - q_1 \neq 0, \text{ grd } (f - gq) < \text{grd } g, \text{ grd } (f - gq_1) < \text{grd } g$$

se tendría (§ 184, teorema 3)

$$q - q_1 \neq 0 \quad \text{y} \quad \text{grd } [f - gq - (f - gq_1)] = \text{grd } [g(q - q_1)] < \text{grd } g$$

lo que es imposible, de donde poniendo $f - gq = r$ e incluyendo el caso en que f es divisible por g:

TEOREMA 9. — *Siendo* K *un cuerpo conmutativo y f y g dos elementos de* K[X], ($g \neq 0$), *existe un par único de polinomios q y r tales que*

(D) $$f = gq + r \quad \text{y} \quad (r = 0 \quad \text{o} \quad \text{grd } r < \text{grd } g).$$

La operación que permite pasar de la pareja (f, g), $g \neq 0$ a la pareja (q, r) verificando (D) se llama la *división euclídea* de los polinomios o también la división de f por g ordenados según las potencias decrecientes de X; q y r se llaman, respectivamente, *cociente y resto en la división euclídea de f por g*; (D) se llama la *igualdad de la división euclídea.* (ATENCIÓN: No olvidar la condición relativa a los grados.)

c) Cálculo efectivo del cociente y del resto

Observemos que si grd $f <$ grd g se tiene

$$f = g0 + f, \quad \text{grd } f < \text{grd } g,$$

luego $q = 0$, $r = f$.

Supongamos, pues, grd $f \geq$ grd g

$$f = a_n X^n + \ldots + a_0 \qquad (a_n \neq 0)$$
$$g = b_m X^m + \ldots + b_0 \qquad (b_m \neq 0) \qquad (m \leq n)$$

y operemos como lo hemos hecho para buscar el cociente de f por g cuando f es divisible por g.

Consideremos el monomio m_0 cociente del monomio dominante de f por el monomio dominante de g sea $m_0 = (a_n/b_m)X^{n-m}$, pongamos

(1) $$f_1 = f - m_0 g.$$

O bien $f_1 = 0$ y f es divisible por g, siendo el cociente m_0 o bien grd $f_1 <$ grd f. Si grd $f_1 <$ grd g la operación está terminada $q = m_0$, $r = f_1$.

Supongamos, pues, $f_1 \neq 0$ y grd $f_1 \geq$ grd g, podemos calcular el monomio m_1 cociente de los monomios dominantes de f_1 y de g, pongamos

(2) $$f_2 = f_1 - m_1 g$$

con $f_2 = 0$ o grd $f_2 <$ grd f_1. Si $f_2 \neq 0$ y grd $f_2 \geq$ grd g podemos recomenzar la operación. Supongamos que por este procedimiento se haya hallado

(h) $$f_h = f_{h-1} - m_{h-1} g.$$

la operación se podrá seguir si $f_h \neq 0$ y grd $f_h \geq$ grd g. Como

$$\text{grd } f_h < \text{grd } f_{h-1} < \ldots < \text{grd } f_1 < \text{grd } f,$$

hay un entero máximo k tal que

$$(k) \qquad f_k = f_{k-1} - m_{k-1}g$$

con

$$\text{no } (f_k \neq 0 \quad \text{y} \quad \text{grd } f_k \geq \text{grd } g) \Leftrightarrow (f_k = 0 \quad \text{o} \quad \text{grd } f_k < \text{grd } g)$$

sumemos miembro a miembro las igualdades de (1) a (k) obtenemos

$$f_k = f - (m_0 + \ldots + m_k)g, \qquad (f_k = 0 \quad \text{o} \quad \text{grd } f_k < \text{grd } g),$$

luego según el teorema 9

$$q = m_0 + \ldots + m_k, \qquad r = f_k.$$

EJEMPLO

Efectuar la división euclídea de f por g

$$f = 3X^5 + X^4 - 6X^2 + 5X - 1, \qquad g = 2X^3 - X + 1$$

$$\begin{array}{rl|l}
f = & 3X^5 + X^4 \qquad\qquad - 6X^2 + 5X - 1 & 2X^3 - X + 1 = g \\ \hline
-m_0g = & -3X^5 \qquad + \frac{3}{2}X^3 - \frac{3}{2}X^2 & \frac{3}{2}X^2 + \frac{1}{2}X + \frac{3}{4} = q \\ \cline{2-2}
f_1 = f - m_0g = & X^4 + \frac{3}{2}X^3 - \frac{15}{2}X^2 + 5X - 1 & \\
-m_1g = & -X^4 \qquad + \frac{1}{2}X^2 - \frac{1}{2}X & \\ \cline{2-2}
f_2 = f_1 - m_1g = & \frac{3}{2}X^3 - 7X^2 + \frac{9}{2}X - 1 & \\
-m_2g = & -X^3 \qquad + \frac{3}{4}X - \frac{3}{4} & \\ \cline{2-2}
r = f_3 = f_2 - m_2g = & -7X^2 + \frac{21}{4}X - \frac{7}{4} &
\end{array}$$

Se puede simplificar la escritura efectuando de memoria el producto m_hg y la sustracción $f_h - m_hg$

$$\begin{array}{rl|l}
f = & 3X^5 + X^4 \qquad\qquad - 6X^2 + 5X - 1 & 2X^3 - X + 1 = g \\ \cline{3-3}
f_1 = & X^4 + \frac{3}{2}X^8 - \frac{15}{2}X^2 + 5X - 1 & \frac{3}{2}X^2 + \frac{1}{2}X + \frac{3}{4} = q \\
f_2 = & \frac{3}{2}X^3 - 7X^2 + \frac{9}{2}X - 1 & \\
r = & -7X^2 + \frac{21}{4}X - \frac{7}{4} &
\end{array}$$

Se procurará dejar los lugares vacíos correspondientes a los monomios con coeficientes nulos.

d) Caso en que K es un anillo conmutativo unitario

Se observará que en el cuerpo K, *las únicas divisiones efectuadas son las divisiones por b_m,* coeficiente dominante de g (en la búsqueda de los monomios m_0, m_1, ..., m_h, ..., m_k). Luego la teórica y la práctica se aplica a todo par de polinomios de K[X], K *anillo conmutativo unitario* con la condición de que el *coeficiente dominante del polinomio divisor g sea inversible en el anillo* **K**: este es el caso si g es unitario.

EJEMPLO

División por $X - a$. En el anillo A[X], A anillo conmutativo unitario, efectuar la división euclídea de

$$f(X) = a_n X^n + a_{n-1} X^{n-1} + \dots + a_0$$

por $X - a$; el resto es nulo o de grado cero, el cociente es de grado $n - 1$, de donde

$$(q_{n-1}, \dots, q_0, r \in A)$$

$$a_n X^n + \dots + a_0 = (q_{n-1} X^{n-1} + \dots + q_0)(X - a) + r$$

con

$$\begin{cases} a_n = q_{n-1} \\ a_{n-1} = q_{n-2} - a q_{n-1} \\ \vdots \\ a_h = q_{h-1} - a q_h \\ \vdots \\ a_1 = q_0 - a q_1 \\ a_0 = r - a q_0 \end{cases} \Rightarrow \begin{cases} q_{n-1} = a_n \\ q_{n-2} = a_{n-1} + a q_{n-1} \\ \vdots \\ q_{h-1} = a_h + a q_h \\ \vdots \\ q_0 = a_1 + a q_1 \\ r = a_0 + a q_0. \end{cases}$$

Por ejemplo, dividir $f(X)$ por $X + 2 = X - (-2)$ con

$$f(X) = 5X^5 - 2X^3 + X^2 + 7X - 3$$

se obtiene

$$q(X) = 5X^4 - 10X^3 + 18X^2 - 35X + 77, \qquad r = -157$$

este método es muy simple por calcular el resto de $f(X)$ por $X - a$, que no es otro que $f(a)$; en efecto, la relación

$$f(X) = q(X)(X - a) + r(X)$$

da, sustituyendo X por a, $r = f(a)$.

EJERCICIOS

1. Efectuar las divisiones euclídeas de f por g en **Q**(X) con

a) $f = 7X^4 - X^3 + 2X - 4, \quad g = 2X^2 - 3X - 5$
b) $f = X^8 - 1, \quad g = X^3 - 1.$

2. Calcular $f(a)$ para los polinomios f de (**Q**[X] o **C**[X]) y los escalares a siguientes

a) $f = 2X^5 - 5X^3 + 8X, \quad a = -3 \in \mathbf{Q}$
b) $f = 4X^3 + X^2, \quad a = -1 - i \in \mathbf{C}.$

3. Hallar el cociente de $(X-2)^{2n}+(X-1)^n-1$ por $(X-1)(X-2)$.

4. Verificar las igualdades siguientes válidas en A[X], anillo conmutativo unitario,

$$(a \in A,\ n \in \mathbf{N})$$

$$X^n - a^n = (X-a)(X^{n-1} + aX^{n-2} + \dots + a^hX^{n-h-1} + \dots + a^{n-1})$$
$$X^n - a^n = (X+a)(X^{n-1} - aX^{n-2} + \dots + (-a)^hX^{n-h-1} + \dots + (-a)^{n-1}) + a^n[(-1)^n + 1].$$
$$X^n + a^n = (X-a)(X^{n-1} + aX^{n-2} + \dots + a^hX^{n-h-1} + \dots + a^{n-1}) + 2a^n$$
$$X^n + a^n = (X+a)(X^{n-1} - aX^{n-2} + \dots + (-a)^hX^{n-h-1} + \dots + (-a)^{n-1}) + a^n[(-1)^n + 1]$$

Hacer explícitas estas fórmulas para $2 \leq n \leq 5$.

5. Hallar el cociente q_0 de $f = X^{n+1} - (n+1)aX^n + na^{n+1}$ por $(X-a)$, después el cociente q de q_0 por $X-a$; deducir una igualdad en A[X] (A anillo conmutativo unitario).

6. ¿Se puede efectuar la división euclídea de f por g con

$$f = 6X^3 + X^2 + 7X \qquad g = 3X^2 + 2X - 1$$

en **Z**[X]?

190. Divisibilidad en el anillo principal K[X] (K cuerpo conmutativo)

a) TEOREMA 10. — *Siendo* K *un cuerpo conmutativo, el anillo* K[X] *es principal.*

Sabemos ya (§ 184, teorema 4) que K[X] es un anillo unitario íntegro. Es suficiente demostrar que todo ideal I de K[X] es principal (§ 94, *c*).

Sea $I \neq \{0\}$, hay en I polinomios no nulos, el conjunto de sus grados admite un elemento mínimo n, sea f un polinomio de I que tiene este grado mínimo.

Consideremos un polinomio arbitrario p de I, efectuemos la división euclídea de p por f

$$p = fq + r, \qquad (r = 0 \quad \text{o} \quad \operatorname{grd} r < \operatorname{grd} f);$$

ahora bien, $r = p - fq$ pertenece a I; luego, siendo f de grado mínimo en I, es imposible que grd $r <$ grd f, luego $r = 0$ y $p = fq$, luego $I = (f)$.

Si $I \neq \{0\}$ es tal que $I = (f) = (g)$, el teorema 2 (§ 98) y el corolario del teorema 4 (§ 184) muestran que existe $\lambda \neq 0$ de K tal que $g = \lambda f$, luego:

COROLARIO. — *Siendo* K *un cuerpo conmutativo, a todo ideal* $I \neq \{0\}$ *de* K[X] *se puede asociar un polinomio unitario* f *no nulo, único tal que* $I = (f)$.

Se deduce de estos resultados que todo lo que hemos dicho del anillo **Z** en los §§ 99, 100 y 101 del capítulo 7 (m. c. d., m. c. m., divisibilidad) puede trasladarse al anillo K[X] reemplazando ± 1 por un elemento cualquiera de K* Los elementos extremales de K[X] se llaman *polinomios irreducibles* de K[X]: estos son los polinomios f no nulos de grado > 0 (no inversibles) y divisibles solamente por λ y λf ($\lambda \in K^*$).

Vamos a servirnos ahora de algunos resultados de estos párrafos enunciándolos con el vocabulario utilizado en K[X]. En el párrafo siguiente haremos lo mismo para los polinomios irreducibles.

OBSERVACION

La hipótesis K es un cuerpo es esencial para afirmar que K[X] es un anillo principal (ver ej. 2 más abajo).

b) Siendo $f_1, f_2, \ldots, f_n$ polinomios distintos de cero, existe un *polinomio unitario único d y un polinomio unitario único m* tal que

$$(f_1) + \ldots + (f_n) = (d)$$
$$(f_1) \cap \ldots \cap (f_n) = (m)$$

d es el m. c. d. y m el m. c. m. de $f_1, \ldots, f_n$. Diremos, siendo λ un elemento no nulo de K, que λd (resp. λm) es un m. c. d. (resp. un m. c. m.) de $f_1, \ldots, f_n$. Se observará que el calificativo máximo (resp. el mínimo) se aplica al grado de los polinomios considerados.

Se ruega al lector a manera de ejercicio escribir los enunciados de los teoremas análogos a los teoremas 4 y 4', 5 y 5', 6 y 6', 7, 8, 9 y 9' de los §§ 99 al 101, por ejemplo: *Siendo $f_1, f_2, \ldots, f_n$ polinomios distintos de cero, las propiedades siguientes son equivalentes:*

1. *$f_1, \ldots, f_n$ son extraños en su conjunto (o primos entre sí en su conjunto), es decir, sus únicos divisores comunes son los polinomios constantes*

2. *El* m. c. d. *de $f_1, \ldots, f_n$ es el polinomio* 1.

3. *Existen polinomios $u_1, \ldots, u_n$ de* K[X] *tales que*

$$f_1u_1 + \ldots + f_nu_n = 1 \qquad \text{(igualdad de Bezout).}$$

4. *Para todo polinomio g de* K[X] *existen los polinomios $g_1, \ldots, g_n$ de* K[X] *tales que*

$$g = f_1g_1 + \ldots + f_ng_n.$$

Se puede precisar la igualdad de BEZOUT para dos polinomios f y g primos entre sí. Sabemos que existe u y v tales que

$$\text{(B)} \qquad fu + gv = 1$$

efectuemos las divisiones euclídeas respectivas de u por g y v por

$$u = gu_1 + u_0 \qquad (u_0 = 0 \quad \text{o} \quad \text{grd } u_0 < \text{grd } g)$$
$$v = fv_1 + v_0 \qquad (v_0 = 0 \quad \text{o} \quad \text{grd } v_0 < \text{grd } f),$$

de donde

$$fu_0 + gv_0 + fg(u_1 + v_1) = 1$$

si u_0 (resp. v_0) es nulo, g divide u, luego $1 = fu + gv$ y grd $g = 0$ (resp. $f = 0$); supondremos f y g de grado estrictamente positivo, luego u_0 y v_0 son no nulos. Tenemos entonces grd $(fu_0 + gv_0) <$ grd f + grd g, luego si $u_1 + v_1 \neq 0$ la igualdad anterior es imposible, pues grd $[fg(u_1 + v_1)]$ sería estrictamente superior a (grd f + grd g). En consecuencia, $u_1 + v_1 = 0$. Demostremos que el par (u_0, v_0) así determinado (grd $u_0 <$ grd g, grd $v_0 <$ grd f) es único; sean,

en efecto, dos parejas (u_0, v_0), (u_2, v_2) las que verifican estas condiciones sobre el grado

$$fu_0 + gv_0 = fu_2 + gv_2 = 1 \Rightarrow f(u_0 - u_2) = g(v_2 - v_0)$$

f primo con g divide $v_2 - v_0$ (teorema de GAUSS, ver § 99, teorema 7); ahora bien, grd $f >$ grd $(v_2 - v_0)$, luego $v_2 - v_0 = 0$ (§ 189, *a*) y $u_2 = u_0$, luego: *si f y g son dos polinomios primos entre sí de* K[X], *de grado no nulo, existe un único par* (u_0, v_0) *de polinomios* K[X] *verificando la relación*

$$(B_0) \qquad fu_0 + gv_0 = 1, \quad \text{grd } u_0 < \text{grd } g, \quad \text{grd } v_0 < \text{grd } f).$$

c) Algoritmo de Euclides para buscar el m. c. d. de dos polinomios

En el caso de **Z** se sabe factorizar todo entero en producto de números primos: resulta de ello un método efectivo para hallar el m. c. d. (aplicación 2 del teorema 11, § 102). En K[X] demostraremos la existencia de la descomposición de todo polinomio de K[X] en producto de polinomios irreducibles, pero no hay en general procedimiento alguno para calcular efectivamente estos factores irreducibles; no podemos, pues, emplear este método para calcular efectivamente el m. c. d. de dos polinomios; el algoritmo de EUCLIDES (ver § 99, ej. 2) nos permitirá calcular efectivamente el m. c. d. de f y de g.

Sea f y g dos polinomios distintos de cero, efectuemos las divisiones euclídeas siguientes parándonos cuando tengamos un resto parcial igual a cero

$$\begin{array}{ll} f = gq + r & (\text{grd } r < \text{grd } g) \\ g = rq_1 + r_1 & (\text{grd } r_1 < \text{grd } r) \\ \vdots & \vdots \\ r_{n-2} = r_{n-1}q_n + r_n & (\text{grd } r_n < \text{grd } r_{n-1}) \\ r_{n-1} = r_n q_{n+1} & (r_{n+1} = 0) \end{array}$$

se ve fácilmente que los divisores comunes a f y g son idénticos a los divisores comunes a r_{h-1} y r_h; éstos son, pues, los divisores de r_n; *r_n último resto no nulo es, pues, un* m. c. d. Por otra parte, estas igualdades permiten encontrar efectivamente u_h y v_h tales que $r_h = fu_h + bv_h$, luego u y v tales que

$$r_n = fu + gv.$$

La operación puede escribirse de la manera siguiente escribiendo los cocientes encima del divisor correspondiente y recopiando cada resto parcial a su nuevo lugar de divisor

	q	q_1	q_2	q_3	...
f	g	r	r_1	r_2	...
r	r_1	r_2	r_3	r_4	...

EJERCICIOS

1. Si K y L son dos cuerpos conmutativos tales que $K \subset L$, demostrar que dos polinomios con coeficientes en K son distintos en K[X] si y sólo si lo son en L[X].

2. Demostrar que los dos polinomios $X^2 + 1$ y $2X$ de **Z**[X] son primos entre sí. Deducir que el anillo **Z**[X] no es principal (considerar el ideal I engendrado por $X^2 + 1$ y $2X$ que, si fuera principal, sería $(1) = \mathbf{Z}[X]$ lo que sería falso).

¿Qué se puede decir en **Z**[X] del ideal engendrado por $X^2 + 1$ y X?

3. En el anillo **Z**[X] demostrar que el ideal engendrado por $X^2 + 1$ es primo y no maximal.

4. Calcular el m. c. d. de f y g para

a) $f = X^4 + X^3 - 3X^2 - 4X - 1,\qquad g = X^3 + X^2 - X - 1$

b) $f = X^4 - 10X^2 + 1,\qquad g = X^4 - 4\sqrt{2}X^3 + 6X^2 - 4\sqrt{2} + 1$

c) $f = X^5 - iX^4 + X^3 - X^2 + iX - 1,\qquad g = X^4 - iX^3 + 3X^2 - 2iX + 2.$

5. Calcular el m.c.d. de $f = 2X^6 - 5X^5 - 14X^4 + 36X^3 + 86X^2 + 12X - 31,$

$$g = 2X^5 - 9X^4 + 2X^3 + 37X^2 + 10X - 14,\qquad h = X^3 - 2X - 1.$$

6. Siendo f y g primos entre sí hallar todas las parejas u y v tales que $fu + gv = 1$ (ver § 99, ej. 1, *a*).

7. Se consideran los polinomios $f(X)$ tales que $f(X) + 1$ sea divisible por $(X - 1)^3$ y $f(X) + 2$ divisible por X^4, hallar los polinomios $f(X)$ de grado mínimo, seguidamente todos los polinomios $f(X)$.

(Observar que $(X - 1)^3$ y X^4 son distintos y utilizar la fórmula de Bezout (B_0) —con limitación de los grados— y el ejercicio 6 de más arriba.) Ver otro método, § 192, ej. 8.

191. Polinomios irreducibles en K[X] (K cuerpo conmutativo)

Siendo K un cuerpo conmutativo, f es un polinomio *irreducible* de K[X] si no es inversible (es decir, si grd $f > 0$) y si no es divisible más que por λ y λf (λ elemento cualquiera de K*).

Siendo divisible el polinomio cero por cualquier polinomio: *un polinomio irreducible es siempre no nulo.*

Por otra parte, es evidente que *un polinomio de primer grado es siempre irreducible.*

Observemos que si L es un supercuerpo conmutativo de K, f elemento de K[X] es también elemento de L[X]; pero f puede ser irreducible en el anillo K[X] y reducible en L[X]. Por ejemplo,

$$X^2 - 2 = (X - \sqrt{2})(X + \sqrt{2})$$
$$X^2 + 1 = (X - i)(X + i)$$

$X^2 - 2$ y $X^2 + 1$, irreducibles en **Q**[X], son reducibles, respectivamente, en **R**[X] y **C**[X].

Por un abuso de lenguaje en lugar de decir f es un elemento irreducible de K[X] diremos algunas veces *f con coeficientes en* K *es irreducible (o reducible) sobre* K *(o sobre* L *supercuerpo de* K*).*

En el anillo principal K[X], tenemos resultados análogos a los teoremas 10 y 11 del § 102.

Sea f un polinomio de grado no nulo, no irreducible, admite divisores q tales que

$$0 < \text{grd } q < \text{grd } f$$

entre esos divisores q, existen polinomios de grado mínimo. Sea g_1 uno de ellos: es evidente que g_1 es irreducible.

Se tiene, pues,

$$f = g_1 f_1$$

si f_1 no es irreducible, si existe un divisor g_2 irreducible de f_1: $f_1 = g_2 f_2$, de donde

$$f = g_1 g_2 f_2$$

tendremos entonces

$$f = g_1 g_2 \ldots g_h f_h$$

siendo $g_1, g_2, \ldots, g_h$ irreducible y

$$\text{grd } f > \text{grd } f_1 > \ldots > \text{grd } f_h$$

al cabo de un número finito de operaciones obtendremos un polinomio f_k de grado cero o bien irreducible.

En consecuencia, tendremos con $g_1, \ldots, g_n$ irreducibles y unitarios λ elemento de K*

$$f = \lambda g_1 g_2 \ldots g_n$$

se demostraría que esta descomposición es única (salvo el orden), como en el § 102; esto no quiere decir que los polinomios $g_1, \ldots, g_n$ sean distintos, de donde reagrupando los polinomios g_i iguales entre sí:

Teorema 11. — *Todo polinomio f de grado no nulo de* K[X] *admite al menos un factor irreducible y puede escribirse de una manera única (salvo el orden)*

$$f = \lambda(g_1)^{k_1} \ldots (g_m)^{k_m}$$

en donde λ es un elemento de K*, *$g_1, \ldots, g_m$ polinomios unitarios irreducibles sobre* K, *distintos dos a dos y $k_1, \ldots, k_m$ enteros estrictamente positivos.*

Esta descomposición se llama *factorización canónica de f en* K[X]. Este anillo K[X] es, pues, un *anillo factorial* (ver capítulo 5, ej. 106).

EJERCICIOS

1. Si f es un polinomio irreducible, el ideal (f) es maximal; demostrar utilizando la fórmula de Bezout que el *anillo cociente* K[X]/(f) tiene una estructura de cuerpo conmutativo. Demostrar en particular que **R**[X]/($X^2 + 1$) es un cuerpo isomorfo a **C**.

2. Hallar el m. c. d. de dos polinomios descompuestos en factores irreducibles (ver § 102).

192. Raíces de un polinomio de K[X] (K cuerpo conmutativo)

a) DEFINICIÓN 6. — *Siendo f un polinomio de* K[X], *se dice que el elemento α de* K *es raíz de f si $f(\alpha) = 0$.*

Siendo L *un supercuerpo de* K *y f siempre un polinomio de* K[X] *si α de* L *es tal que $f(\alpha) = 0$ diremos que α es una raíz de f en* L.

Se ve en los siguientes ejemplos que una raíz α de f en L puede que no sea raíz de f en K

$$f(X) = X^2 - 2 \in \mathbf{Q}[X], \quad f(\sqrt{2}) = 0, \quad \sqrt{2} \in \mathbf{R}, \quad \sqrt{2} \notin \mathbf{Q}$$
$$f(X) = X^2 + 1 \in \mathbf{Q}[X], \quad f(i) = 0, \quad i \in \mathbf{C}, \quad i \notin \mathbf{Q}.$$

Cuando digamos que α es raíz de $f \in K[X]$, sobrentenderemos que α es un elemento de K.

TEOREMA 12. — *α elemento de* K *es raíz del polinomio f de* K[X] *si y sólo si f es divisible por $X - \alpha$.*

La igualdad de la división de f por $X - \alpha$

$$f(X) = (X - \alpha)g(X) + h(X) \qquad (h = 0 \quad \text{o} \quad \text{grd } h = 0),$$

luego h pertenece a K, de donde sustituyendo X por α, $f(\alpha) = h$.

Este teorema nos dice que todo polinomio de K[X] que tiene una raíz en K no es irreducible; *la recíproca no es cierta*: así $(X^2 + 1)^2$ reducible sobre **R** está desprovisto de raíz en **R**.

Siendo α raíz de f, puede que f sea divisible no solamente por $X - \alpha$, sino por $(X - \alpha)^h$ $(h > 1)$; hay, pues, $k > 0$ máximo tal que f sea divisible por $(X - \alpha)^k$, de donde la definición:

DEFINICIÓN 7. — *Se llama* orden de multiplicidad *de una raíz α de un polinomio f el entero mayor k tal que f sea divisible por $(X - \alpha)^k$.*

Si $k = 1$ se dice que α es raíz simple de f.

Si $k > 1$ se dice que α es raíz múltiple de orden k de f.

Luego si α es raíz de orden k de f

$$f(X) = (X - \alpha)^k g(X), \qquad g(\alpha) \neq 0$$

pues $g(\alpha) = 0$ implicaría la divisibilidad de g por $X - \alpha$, luego la de f por $(X - \alpha)^{k'}$ con $k' > k$.

Algunas veces se extiende esta definición a α no raíz de f diciendo que α, tal que $f(\alpha) \neq 0$, es raíz de orden cero de f.

EJERCICIOS

1. Siendo L un supercuerpo conmutativo de K, α una raíz de orden h en K del polinomio f de K[X], demostrar que el orden de α raíz en L de f considerado como polinomio de L[X] es también h.

2. Dos polinomios f y g de K[X] primos entre sí no tienen naturalmente raíces comunes en K, demostrar que tampoco tienen en todo supercuerpo conmutativo de K (utilizar la igualdad de BEZOUT).

3. ¿Se pueden extender las definiciones y resultados precedentes al caso en que K es sólo un anillo conmutativo unitario? (Cuidado con los divisores de cero. Se demostrará que el polinomio unitario $X-\alpha$ no es divisor de cero en K[X], K anillo conmutativo unitario.)

b) Sean $\alpha_1, \ldots, \alpha_m$, m raíces distintas de orden respectivo $k_1, \ldots, k_m$ del polinomio f de K[X], se ve fácilmente que $(X-\alpha_1)^{k_1}, \ldots, (X-\alpha_m)^{k_m}$ son extraños *dos a dos* y, en consecuencia, que f es divisible por su producto (ver resultados análogos aquellos de los corolarios del teorema 7, § 99), luego:

TEOREMA 13. — *Si los elementos $\alpha_1, \ldots, \alpha_m$ de* K *son raíces dos a dos distintas, de orden respectivo $k_1, \ldots, k_m$ (estrictamente positivos) del polinomio f de* K[X], *existe un polinomio g de* K[X] *tal que*

$$f(X) = (X-\alpha_1)^{k_1} \ldots (X-\alpha_m)^{k_m} g(X). \tag{1}$$

Por otra parte, $f \neq 0$ si y sólo si $g \neq 0$ y

$$\text{grd. } f-(k_1+\ldots+k_m) = \text{grd } g \geq 0,$$

luego si se considera *todas las raíces* de un polinomio no nulo, este conjunto es finito y la suma de su orden de multiplicidad es inferior al grado de f. Resulta que un polinomio f de *grado n a lo sumo*(36) y que tiene al menos $n+1$ raíces es nulo:

COROLARIO 1. — *Si f es un polinomio de* K[X] *de grado n a lo sumo:*

1. *Si $f \neq 0$ la suma de los órdenes de multiplicidad de las raíces de f es inferior o igual a n.*
2. *Si la suma de los órdenes de multiplicidad de las raíces es estrictamente superior a n, $f = 0$.*

Sean entonces dos polinomios f y g de grado n a lo sumo de K[X], tales que las funciones asociadas $\tilde{f}$ y $\tilde{g}$ toman el mismo valor para $n+1$ valores distintos de la variable, se tiene entonces $f-g=0$; en consecuencia, $f=g$, de donde:

COROLARIO 2. — *Si $\tilde{f}$ y $\tilde{g}$ son las funciones polinomios asociadas, respectivamente, a dos polinomios f y g de grado n a lo sumo, de* K[X], *si $\tilde{f}$ y $\tilde{g}$ toman el mismo valor para $n+1$ valores distintos de la variable, $f = g$.*

Consideremos el homomorfismo φ de K[X] en $\mathfrak{F}$(K, K) (§ 185)

$$f \to \varphi(f) = \tilde{f}.$$

(36) Recordemos que decimos, por abuso de lenguaje, «f es de grado n a lo sumo» en lugar de «$f \neq 0$ o grd $f \leq n$».

Sean f y g dos polinomios tales que $\tilde{f} = \tilde{g}$, es decir, tales que

$$(\forall x \in K) \qquad \tilde{f}(x) = \tilde{g}(x)$$

si K es infinito estas funciones serán iguales para $n+1$ valores distintos de la variable, siendo $n = \sup$ (grd f, grd g) y $f = g$; luego el homomorfismo estudiado es inyectivo, también biyectivo, puesto que sabemos que es suprayectivo.

Si K es finito los ejemplos (ver § 185, ej. 1) muestran que este homomorfismo no es inyectivo. Relacionando estos resultados con los resultados ya obtenidos en el § 185 podemos enunciar:

TEOREMA 14. — *Si* K *es un cuerpo conmutativo la aplicación* $f \to \tilde{f}$ *de* K[X] *en* $\mathcal{F}$(K, K) *es un isomorfismo si y sólo si* K *es infinito.*

Por consiguiente, *en este caso* (K *infinito*) *no hay inconveniente en representar por la misma letra f el polinomio y la función polinómica asociada*: es lo que haremos en particular cuando K es **Q**, **R** o **C**.

EJERCICIOS

4. Demostrar el teorema 13 sin utilizar las propiedades del anillo principal K[X] razonando por inducción y utilizando únicamente el hecho de que K es un anillo íntegro.

Extender los corolarios y el teorema 14 al anillo A[X] y al anillo $\mathcal{F}$(A, A) (siendo A un anillo íntegro unitario).

5. Si $\alpha_1, \ldots, \alpha_n$ son n valores *distintos* de un cuerpo conmutativo K, demostrar que existe un polinomio y sólo uno f de K[X] tal que

$$\text{grd } f \leq n-1 \quad \text{y} \quad (i = 1, 2, \ldots, n) \qquad f(\alpha_i) = \beta_i$$

donde $\beta_1, \ldots, \beta_n$ son n valores cualesquiera de K. Demostrar que este polinomio es

$$f(X) = \sum_{i=1}^{n} \beta_i \frac{(X-\alpha_1)\ldots(X-\alpha_{i-1})(X-\alpha_{i+1})\ldots(X-\alpha_n)}{(\alpha_i-\alpha_1)(\alpha_i-\alpha_{i-1})(\alpha_i-\alpha_{i+1})(\alpha_i-\alpha_n)}$$

(fórmula de interpolarización de LAGRANGE). Ver otra demostración § capítulo 12, ej. 379.

c) Caracterización de las raíces múltiples de un polinomio de K[X], donde K es un cuerpo conmutativo de característica nula

Sea $f \neq 0$ de grado n, y α una raíz de orden $k > 0$, se tiene $k \leq n$ y

$$f(X) = (X-\alpha)^k g(X), \qquad g(\alpha) \neq 0$$

de donde

$$f'(X) = (X-\alpha)^{k-1}[kg(X) + (X-\alpha)g'(X)] = (X-\alpha)^{k-1}h(X)$$

con $h(\alpha) = kg(\alpha) \neq 0$, luego α es una raíz de orden $k-1$ de f'. Recíprocamente si una raíz α de f es una raíz de orden $k-1$ de f' es raíz de orden k de f. Supongamos raíces de orden $h > 0$ de f, se tendrá $h-1 = k-1$, de donde $h = k$.

En consecuencia, si α es una raíz de orden k de f, α será raíz de orden $k-1$ de f' y de manera más general raíz de orden $k-i$ de $f^{(i)}$ para $i<k$ y no será raíz de $f^{(k)}$.

Recíprocamente sea α una raíz de f de grado n, tal que

$$f(\alpha)=f'(\alpha)\ldots=f^{(k-1)}(\alpha)=0, \qquad f^{(k)}(\alpha)\neq 0$$

desde luego $k\leq n$; la fórmula de TAYLOR (§ 188)

$$f(X)=f(\alpha)+(X-\alpha)f'(\alpha)+\ldots+(X-\alpha)^n\frac{f^{(n)}(\alpha)}{n!}$$

nos da

$$f(X)=(X-\alpha)^k g(X), \qquad g(\alpha)=\frac{f^{(k)}(\alpha)}{k!}\neq 0$$

pues

$$g(X)=1\frac{f^{(k)}(\alpha)}{k!}+(X-\alpha)\frac{f^{(k+1)}(\alpha)}{k+1}+\ldots+(X-\alpha)^{n-k}\frac{f^{(n)}(\alpha)}{n!}$$

luego α es raíz de orden k de f, de donde:

TEOREMA 15. — *Siendo f un elemento de* K[X], K *cuerpo conmutativo de característica nula, las tres propiedades siguientes son equivalentes:*

1. *α es raíz de orden k de f.*
2. *α es raíz de f y es raíz de orden $k-1$ de f'.*
3. $f(\alpha)=f'(\alpha)=\ldots=f^{(k-1)}(\alpha)=0$, $f^{(k)}(\alpha)\neq 0$.

OBSERVACION

La segunda condición, en general, es más práctica, ya que en ella no interviene el orden de multiplicidad de α para los polinomios $f, f', \ldots, f^{(k-1)}$.

EJERCICIOS

6. Hallar la relación entre p y q de **C** para que X^3+pX+q tenga una raíz doble α, ¿cuál es esta raíz doble?
7. Hallar los polinomios f de grado n de K[X] (K cuerpo conmutativo de característica nula) divisibles por f'.
8. Resolver el ejercicio 7 del § 190 con la ayuda de los polinomios derivados.
9. En qué se transforma el teorema 15 si K es de característica p. Dar ejemplos.

193. Polinomios irreducibles de C[X] y R[X]. Aplicaciones

a) Cuerpo algebraicamente cerrado

Sea un cuerpo conmutativo K tal que todo polinomio de K[X] de grado no nulo tenga al menos una raíz α_1 en K, se tendrá (§ 192, teorema 12)

$$f(X)=(X-\alpha_1)f_1(X) \qquad (\text{grd } f_1=\text{grd } f-1)$$

se ve, pues, por inducción que si grd $f = n > 0$

$$f(X) = (X - \alpha_1) \dots (X - \alpha_n) f_n(X) \qquad (\text{grd } f_n = 0)$$

con $f_n = a_n$, coeficiente dominante de f.

Siendo irreducible todo polinomio de primer grado, resulta que en un tal cuerpo K, los únicos polinomios irreducibles son los polinomios del primer grado, de donde:

Teorema y definición 8. — *Para un cuerpo conmutativo* K *las tres propiedades siguientes son equivalentes:*

1. *Todo polinomio de* K[X] *de grado no nulo admite al menos una raíz en* **K.**
2. *Todo polinomio* f *de* K[X] *de grado* $n > 0$ *admite* n *raíces* $\alpha_1, \dots, \alpha_n$ en K.
3. *Los únicos polinomios irreducibles de* K[X] *son los polinomios de primer grado.*

Todo cuerpo poseyendo una de las propiedades precedentes se llama cuerpo algebraicamente cerrado.

Así **Q** y **R** no son algebraicamente cerrados, pues $X^2 - 2$ no tiene raíz en **Q** y $X^2 + 1$ no tiene raíz en **R**.

Dado un polinomio irreducible f de grado n de K[X], siendo K no algebraicamente cerrado, se demuestra que existe un supercuerpo conmutativo minimal Δ de K, único, salvo un isomorfismo, tal que f tenga raíces en Δ, Δ se llama el *cuerpo de las raíces de* f (ver ej. 348 al 350). Se demuestra un resultado mucho más interesante, el teorema de Steinitz: todo cuerpo conmutativo K puede ser encajado(*) en un supercuerpo conmutativo, algebraicamente cerrado Ω, único, salvo un isomorfismo, a Ω se le llama *cierre algebraico de* K. La demostración de este teorema es de un nivel superior al de este curso.

b) Caso del cuerpo C

La invención de los números complejos (llamados entonces números imaginarios) en el siglo XVI, seguida rápidamente por la resolución de las ecuaciones de tercer y cuarto grado con coeficientes en **C**, que tienen, respectivamente, 3 y 4 raíces en **C** (ver capítulo 13, ej. 441 y 442), junto con el hecho de que la ecuación $x^n = a$, tiene n raíces en **C**, hizo presentir desde el siglo XVII que toda ecuación de n-ésimo grado con coeficientes en **C** tenía n raíces en **C** (naturalmente teniendo en cuenta su orden de multiplicidad). Varios matemáticos del siglo XVIII y en particular d'Alembert creyeron demostrar este resultado de varias maneras, pero ninguna era válida; es Gauss quien tiene el mérito de encontrar esta demostración (de hecho al final del siglo XVIII y al principio del siglo XIX, Gauss dio cuatro demostraciones distintas). Admitiremos este teorema:

Teorema de d'Alembert-Gauss. — *El cuerpo* **C** *de los números complejos es algebraicamente cerrado.*

(*) *N. del T.* — Hemos preferido utilizar la palabra «encajado» para traducir «plongé», equivalente a la palabra inglesa «imbedded», por responder mejor al sentido matemático que la de sumergido.

Una de las características de este teorema es de no ser un teorema puramente algebraico; para demostrarlo hay que recurrir no solamente a las *propiedades algebraicas* de **C**, sino también a las *propiedades topológicas* que se derivan de las de **R**.

Las distintas demostraciones se diferencian según la "dosis" de Análisis empleado para demostrarlo: *Es necesario al menos utilizar el hecho de que todo número real estrictamente positivo tiene dos raíces cuadradas en* **R** (ver § 111) *y el hecho de que todo polinomio de grado impar de* **R**[X] *tiene al menos una raíz real* (ver curso de Análisis). Una de estas demostraciones se propone en el ej. 408, final del capítulo 13.

El teorema de D'ALEMBERT-GAUSS y la definición de un cuerpo cerrado demuestra que para todo polinomio $f(X) = a_0 + \dots + a_n X^n$ de **C**[X] con $a_n \neq 0$ se tiene

$$f(X) = a_n(X - \alpha_1) \dots (X - \alpha_n)$$

donde $\alpha_1, \dots, \alpha_n$ son números complejos distintos o no.

Luego si f admite m raíces distintas $\alpha_1, \dots, \alpha_m$ de órdenes respectivos $k_1, \dots, k_m$ se tendrá

$$(1) \qquad f(X) = a_n(X - \alpha_1)^{k_1} \dots (X - \alpha_m)^{k_m}$$

con $k_1 + \dots + k_m = n$, siendo a_n el coeficiente dominante de f. Desarrollaremos las consecuencias de este resultado en el capítulo 13 (ecuaciones algebraicas de coeficiente en **C**).

Observemos, sin embargo, seguidamente que si se pone

$$\bar{f}(X) = \bar{a}_0 + \bar{a}_1 X + \dots + \bar{a}_n X^n$$

el polinomio $\bar{f}$ llamado *polinomio conjugado* del polinomio f de **C**[X], se tiene ($\lambda \in \mathbf{C}$)

$$\overline{(f+g)} = \bar{f} + \bar{g}, \qquad \overline{fg} = \bar{f}\bar{g}, \qquad \overline{(\lambda f)} = \bar{\lambda}\bar{f}.$$

La aplicación $f \to \bar{f}$ de **C**[X] en **C**[X] al ser evidentemente biyectiva las dos primeras igualdades demuestran que esta aplicación es un *automorfismo del anillo* **C**[X]. Pero la última de estas tres igualdades demuestra que esta aplicación no es un automorfismo del espacio vectorial **C**[X] ni del álgebra **C**[X].

Además, esta aplicación es tal que $\overline{(\bar{f})} = f$; luego f y g siendo dos polinomios de **C**[X] la segunda relación anterior demuestra que

$$f \text{ divisible por } g \Leftrightarrow \bar{f} \text{ divisible por } \bar{g}.$$

Finalmente se tiene

$$[f \in \mathbf{C}[X] \quad \text{y} \quad \alpha \in \mathbf{C}] \Rightarrow \overline{f(\alpha)} = \bar{f}(\bar{\alpha}).$$

c) **Caso del cuerpo R**

f pertenece a **R**[X] si y sólo si $f = \bar{f}$, luego si $\alpha \in \mathbf{C}$, f de **R**[X] es divisible en **C**[X] por $(X - \alpha)^k$, k entero estrictamente positivo si y sólo si $f = \bar{f}$ es divisible por $(X - \bar{\alpha})^k$, de donde utilizando la definición 7 del § 192:

TEOREMA 16. — *Si f es un elemento de* $\mathbf{R}[X]$, *el número complejo α es raíz de orden k de f si y sólo si $\bar{\alpha}$ es raíz de orden k de f.*

Sea f un polinomio de $\mathbf{R}[X]$ de grado n, según el teorema de D'ALEMBERT-GAUSS, hay n raíces en $\mathbf{C}$; algunas son reales $\alpha_1, \ldots, \alpha_p$, de órdenes respectivos $h_1, \ldots, h_p$. Las raíces complejas no reales se clasifican por parejas de raíces complejas conjugadas β_1 y $\bar{\beta}_1, \ldots, \beta_q$ y $\bar{\beta}_q$ de órdenes respectivos $k_1, \ldots, k_q$.

Observemos que si $\beta = u + iv$ $(u, v \in \mathbf{R}, v \neq 0)$ se tiene

$$(X - \beta)(X - \bar{\beta}) = X^2 - (\beta + \bar{\beta})X + \beta\bar{\beta} = X^2 - 2uX + u^2 + v^2 = (X - u)^2 + v^2$$

este polinomio es irreducible, pues sus divisores propios siendo de primer grado no pueden ser más que $X - \beta$ y $X - \bar{\beta}$, que no pertenecen a $\mathbf{R}[X]$; luego si f es de grado n, y de coeficiente dominante a_n, tendremos con las notaciones anteriores y esto de modo único, salvo el orden,

$$f(X) = a_n \prod_{i=1}^{p} (X - \alpha_i)^{h_i} \prod_{j=1}^{q} [(X - u_j)^2 + v_j^2]^{k_j} \qquad (2)$$

donde $\alpha_1, \ldots, \alpha_p, u_1, \ldots, u_q, v_1, \ldots, v_q$ son reales ($v_1, \ldots, v_q$ no nulos) y los enteros $h_1, \ldots, h_p, k_1, \ldots, k_q$ verificando

$$h_1 + \ldots + h_p + 2(k_1 + \ldots + k_q) = n = \text{grd } f.$$

Se llama algunas veces "*descomposición de Gauss*" esta factorización de un polinomio real.

Resulta de la fórmula (2) que los únicos polinòmios irreducibles de $\mathbf{R}[X]$ son los polinomios de primer grado y los polinomios de segundo grado con raíces complejas no reales, es decir, los polinomios $aX^2 + bX + c$ tales que $b^2 - 4ac < 0$ (ver § 114, b), de donde:

COROLARIO. — *Los únicos polinomios irreducibles de* $\mathbf{R}[X]$ *son los polinomios de primer grado y los polinomios $aX^2 + bX + c$ tales que $b^2 - 4ac < 0$. Todo polinomio f de* $\mathbf{R}[X]$ *de grado n y de coeficiente dominante a_n se escribe de un modo único, salvo el orden*

$$f(X) = a_n P_1^{h_1} \ldots P_p^{h_p} S_1^{k_1} \ldots S_q^{k_q}$$

Siendo P_i $(1 \leq i \leq p)$ y Q_j $(1 \leq j \leq q)$ los polinomios irreducibles unitarios, respectivamente, de primero y segundo grado, los enteros $h_1, \ldots, h_p, k_1, \ldots, k_q$ verifican la relación

$$h_1 + \ldots + h_p + 2(k_1 + \ldots + k_q) = n.$$

EJERCICIOS

1. Demostrar que (ver § 120, ej. 4 y 5)

$$X^{2p} - 1 = (X - 1)(X + 1) \prod_{k=1}^{p-1} \left(X^2 - 2X \cos \frac{k\pi}{p} + 1 \right)$$

$$X^{2p+1} - 1 = (X - 1) \prod_{h=1}^{h=p} \left(X^2 - 2X \cos \frac{2h\pi}{2p+1} + 1 \right).$$

2. Hallar la descomposición de GAUSS de $X^4 - 2pX^2 + q$ $(p,\ q \in \mathbf{R})$. En el caso en que las raíces de $Y^2 - 2pY + q$ son complejas no reales, es decir, $p^2 - q < 0$, se observará que $q > 0$ y $p + \sqrt{q} > 0$ y se escribirá

$$X^4 - 2pX^2 + q = (X^2 + \sqrt{q})^2 - 2(p + \sqrt{q})X^2$$

aplicación a $X^4 + 1$, $X^4 + X^2 + 1$.

3. Hallar la descomposición de GAUSS de $X^8 + 1$.

194. División de polinomios ordenados según las potencias crecientes

Siendo K siempre un cuerpo conmutativo, consideremos los dos polinomios de K[X] ordenados según las potencias crecientes

$$f(X) = a_m X^m + \dots + a_n X^n \qquad (a_m a_n \neq 0,\ m < n)$$
$$g(X) = b_p X^p + \dots + b_q X^q \qquad (b_p b_q \neq 0,\ p < q).$$

Podemos tratar de aplicar un procedimiento análogo al de la división euclídea de los polinomios, pero empezando por los monomios de menor grado; si es posible, es decir, si $m \geq p$, formemos

$$m_0 = \frac{a_m}{b_p} X^{m-p}, \qquad f_1 = f - m_0 g.$$

Vemos que el grado del monomio de menor grado de f_1 es estrictamente superior al del monomio de menor grado de f.

Es decir, en este principio de operación nos interesamos no del grado del término de mayor grado de un polinomio (es decir, a su grado), sino del grado de su monomio de menor grado, lo que justifica la siguiente definición:

a) Orden de un polinomio

DEFINICIÓN. — *Dado el polinomio* $f = \sum_h a_h X^h$ *de* K[X], *se llama* orden *de* f, *que se representa* $\omega(f)$, *al menor entero* m *tal que*

$$a_m \neq 0.$$

Resulta de esta definición que el *polinomio cero no tiene orden* y que para $f \neq 0$

$$\omega(f) \leq \operatorname{grd} f$$

la igualdad tiene lugar si y sólo si f es un monomio.

Las propiedades de la adición y de la multiplicación de los polinomios de K[X], demuestran que f, g, $f + g$ siendo no nulos

$$\omega(f + g) \geq \inf\ [\omega(f),\ \omega(g)]$$
$$\omega(fg) = \omega(f) + \omega(g).$$

EJERCICIOS

1. ¿En qué caso se tiene $\omega(f+g) = \inf\ [\omega(f), \omega(g)]$?

2. Se reemplaza K por un anillo conmutativo unitario cualquiera, ¿en qué se convierte la relación que da $\omega(fg)$?

b) División según las potencias crecientes

La operación tratada anteriormente podrá iniciarse si $\omega(f) \geq \omega(g)$. Este será el caso cuando $\omega(g) = 0$, *lo que ahora supondremos.* Sea entonces m_0 el cociente del monomio de menor grado de f, es decir, $a_m X^m$, por $b_0 \neq 0$, monomio de menor grado de g, pongamos

$$(1) \qquad f_1 = f - m_0 g$$

tendremos

$$f_1 = (a_m X^m + \dots + a_n X^n) - \left(\frac{a_m}{b_0} X^m \right) (b_0 + b_1 X + \dots + b_q X^q)$$

en consecuencia: o bien $f_1 =$, o bien

$$\omega(f) > \omega(f_1)$$

podemos *empezar nuevamente* la operación sobre f_1 y g y así sucesivamente mientras que $f_1, f_2, \dots, f_{h-1}$ sean no nulos, pues las fórmulas

$$(2) \qquad f_2 = f_1 - m_1 g$$
$$(h+1) \qquad f_{h+1} = f_h - m_h g,$$

en donde m_h es el cociente del monomio de menor grado de f_{h-1} por b_0

$$\omega(f) < \omega(f_1) < \dots < \omega(f_h) < \omega(f_{h+1})$$
$$\text{grd}\ (m_0) = \omega(f),\ \text{grd}\ (m_1) = \omega(f_1),\ \dots,\ \text{grd}\ (m_h) = \omega(f_h),$$

en consecuencia: *o bien encontramos un resto parcial* $f_{k+1} = 0$ *o bien la operación se podrá continuar indefinidamente.*

En el primer caso existe k tal que $f_k \neq 0$, $f_{k+1} = 0$, por adición de las igualdades (h) de 1 a $k+1$, tendremos

$$0 = f - (m_0 + \dots + m_k) g$$

esto se producirá si y sólo si f *es divisible por* g, entonces $q = m_0 + \dots + m_k$ es el *cociente* de f por g: la única diferencia con la división euclídea de f por g es que q está ordenado según las potencias crecientes de X.

En el segundo caso *para todo* h, $f_h \neq 0$, la operación puede seguirse indefinidamente: es lo que sucederá en realidad si f no es divisible por g. Tenemos

$$\omega(f) = \text{grd}\ (m_0) < \text{grd}\ (m_1) < \dots < \text{grd}\ (m_h) < \text{grd}\ (m_{h+1}),$$

luego, cualquiera que sea el entero natural $l \geq \omega(f)$, existe un entero único k tal que

$$\text{grd}\ (m_k) \leq l \qquad \text{y} \qquad \text{grd}\ (m_{k+1}) > l$$

sumando las igualdades (h) de 1 a $k+1$ obtenemos

$$f_{k+1} = f - (m_0 + \dots + m_k)g$$

con

$$\omega(f_{k+1}) = \text{grd } (m_{k+1}) > l$$

existe, pues, dos polinomios q y s definidos por

$$q = m_0 + \dots + m_k, \qquad f_{k+1} = X^{l+1}s$$

tales que

$$f = gq + X^{l+1}s \qquad (\text{grd } q \leq l).$$

Veamos que fijado el entero l, la pareja (q, s) verificando la relación precedente es única. En efecto, la relación

$$f = gq + X^{l+1}s = gq_1 + X^{l+1}s_1 \qquad (\text{grd } q \leq l, \quad \text{grd } q_1 \leq l)$$

da

$$g(q - q_1) = X^{l+1}(s_1 - s)$$

siendo g no nulo y el anillo K[X] íntegro, $q \neq q_1$ es equivalente a $s \neq s_1$. Luego si $q \neq q_1$

$$\omega[g(q - q_1)] = \omega(q - q_1) \leq \text{grd } (q - q_1) \leq l$$
$$\omega[X^{l+1}(s_1 - s)] \geq \omega(X^{l+1}) = l + 1$$

luego la igualdad precedente es imposible con $q \neq q_1$.

Finalmente si $l < \omega(f)$, es decir, $l + 1 \leq \omega(f)$, existe s único tal que $f = X^{l+1}s$ y la igualdad precedente tiene lugar con $q = 0$. De donde:

TEOREMA 17. — *Dados dos polinomios no nulos f y g de* K[X] (K *cuerpo conmutativo*) *verificando $\omega(g) = 0$ y un entero natural cualquiera l existe una pareja única (q, s) de polinomios de* K[X] *verificando*

(C) $$f = gq + X^{l+1}s, \qquad (q = 0 \quad \text{o} \quad \text{grd } (q) \leq l).$$

La determinación de la pareja (q, s) partiendo de la pareja (f, g) se llama *división según las potencias crecientes*, q y $X^{l+1}s$ son, respectivamente, el *cociente y el resto relativos al entero l en esta división.*

Observemos que si f es divisible por g, la relación (C) anterior es siempre válida para todo l, pero el polinomio q en esta relación no es en general el cociente de f por g, no lo será más que si $l \geq \text{grd } f - \text{grd } g$ y se tendrá entonces $f = gq$, con $s = 0$.

La operación se dispone de la manera siguiente: dividamos $1 - X - X^3 - X^6$ por $1 - X - X^2$ con $l = 3$

$$\begin{array}{rl|l}
f = & 1 - X \qquad - X^3 \qquad - X^6 & 1 - X - X^2 = g \\
-m_0 g = & -1 + X + X^2 & 1 + X^2 \qquad = q \\
\hline
f_1 = & X^2 - X^3 \qquad - X^6 & \\
-m_1 g = & -X^2 + X^3 + X^4 & \\
\hline
f_2 = X^4 s & X^4 - X^6 &
\end{array}$$

Se tiene entonces

$$1 - X - X^3 - X^6 = (1 - X - X^2)(1 + X^2) + X^4(1 - X^2).$$

OBSERVACIONES

1. Hemos seguido en la exposición de la división según las *potencias crecientes* un camino diferente al de la exposición de la división euclídea, pues la necesidad de aumentar el grado de q para obtener la unicidad de q y r sólo se percibe cuando se hace efectivamente la división siguiendo las potencias crecientes. Pero se hubiera podido seguir un procedimiento mucho más dogmático, análogo al de la división euclídea (ver ej. 3 más abajo).

2. Si $\omega(g) = m > 0$ para que la operación pueda empezar es necesario suponer $\omega(f) \geq m$. Supongamos que se cumple esta condición y pongamos

$$f = f_1X^m, \qquad g = g_1X^m, \qquad (\omega(g_1) = 0);$$

existe, pues, g y s tal que

$$f_1 = g_1q + X^{l+1}s \qquad (q = 0 \quad \text{o} \quad \text{grd } q \leq l)$$

q y s verifican, pues,

$$f = gq + X^{m+l+1}s \qquad (q = 0 \quad \text{o} \quad \text{grd } q \leq l).$$

3. Si sólo interesa la determinación efectiva de q (grd $q \leq l$) basta efectuar la operación sobre los polinomios f_1 y g_1 tales que

$$f = f_1 + X^{l+1}f_2, \qquad g = g_1 + X^{l+1}g_2$$

$\omega(g) = \omega(g_1) = 0$ implica la existencia de q y s_1 tales que

$$f_1 = g_1q + X^{l+1}s_1 \qquad (q = 0 \quad \text{o} \quad \text{grd } q \leq l)$$

de donde

$$f = gq + X^{l+1}(s_1 + f_2 - g_2q)$$

el cociente de f por g relativo a l es el mismo que el de f_1 por g_1 relativo a l. Esta observación se utiliza en Análisis para buscar un desarrollo limitado de orden l de una fracción racional.

4. Si f es divisible por g encontraremos, después de un número finito de operaciones, un resto parcial $f_k = 0$. En este caso grd q = grd f — grd g, luego si se obtiene en un momento dado un monomio q_h de grado estrictamente superior a (grd f — grd g), se está seguro de que f no es divisible por g.

EJERCICIOS

3. Dado f y g de K[X], con $\omega(g) = 0$ y f no divisible por g, demostrar que existe un polinomio q y uno sólo tal que, cualquiera que sea $l \geq \omega(f)$, se tenga

$$\text{grd } q \leq l < \omega(f - gq)$$

(considerar el conjunto de los polinomios $f - gp$ en que p describe el conjunto de los polinomios de grado $\leq l$. Demostrar que $\omega(f - gp)$ está acotado superiormente, seguidamente que si q hace en $\omega(f - gq) < l$ y deducir de ello que q es único).

4. Dividir 1 por $1 - X$, o 1 por $1 - aX$ en K[X] ($a \in K$), escribir las igualdades relativas al entero n.

5. Dividir $2 - X + 3X^4 - X^5$ por $1 - X + X^3$ según las potencias crecientes con grd $q \leq 7$.

6. Dividir $1 - X \cos a$ por $1 - 2X \cos a + X^2$ ($a \in \mathbf{R}$) según las potencias crecientes, con grd $q \leq n$, el mismo problema con $X \operatorname{sen} a$ y $1 - 2X \cos a + X^2$.

Deducir las fórmulas obtenidas al final del § 121 (sustituir X por 1 en las igualdades obtenidas).

III. Estudio de $K[X_1, ..., X_m]$ (K cuerpo conmutativo)

Representaremos, respectivamente, por 0 y 1 el cero y la unidad de K. Algunos de los resultados demostrados son válidos en casos más generales, lo señalaremos eventualmente en la teoría o en los ejercicios.

195. Anillo $K[X_1, ..., X_m]$. Nociones sobre la divisibilidad

Sabemos ya que $K[X_1, ..., X_m]$ es un *anillo unitario, íntegro.* Según lo que hemos visto en el § 186, $K[X_1, ..., X_m] = K[X_1, ..., X_{m-1}][X_m]$ tiene por elementos inversibles $K[X_1, ..., X_{m-1}]$, luego razonando por recurrencia se ve que: *los elementos inversibles de* $K[X_1, ..., X_m]$ *son los elementos inversibles de* K, *es decir, los elementos no nulos de* K (siendo K un cuerpo).

Un polinomio f de $K[X_1, ..., X_m]$ es *irreducible* si no es inversible y si sólo admite como divisores λ y λf ($\lambda \in K^*$). Luego si f es reducible existen polinomios f_1 y f_2 tales que

$$f = f_1 f_2 \quad \text{y} \quad (0 < \text{grd } f_1 < \text{grd } f).$$

Como en el caso de una indeterminada si L es un supercuerpo conmutativo de K, f polinomio irreducible de $K[X_1, ..., X_m]$ puede ser reducible en $L[X_1, ..., X_m]$; por ejemplo, $X^2 + Y^2$ es un polinomio irreducible de $\mathbf{R}[X, Y]$, pero en $\mathbf{C}[X, Y]$

$$X^2 + Y^2 = (X - iY)(X + iY).$$

Dos elementos f y g de $K[X_1, ..., X_m]$ son *extraños (o también primos entre sí)* si sólo tienen como divisores comunes los polinomios constantes (elementos de K^*). Se definirá igualmente una familia (f_i) $(1 \leq i \leq p)$ de polinomios *extraños en su conjunto* y una familia de *polinomios extraños dos a dos.*

Las definiciones y propiedades precedentes son análogas a las de $K[X]$, ya que son comunes a todos los anillos unitarios íntegros. Pero esta analogía termina pronto, pues para $m \geq 2$ el *anillo* $K[X_1, ..., X_m]$ *no es principal.* Consideremos, en efecto, $K[X, Y]$: X e Y son manifiestamente distintas, luego si $K[X, Y]$ fuera un anillo principal como $\mathbf{Z}$ y como $K[X]$, existiría u y v de $K[X, Y]$ tales que (ver capítulo 5, § 99 y ej. 105)

$$u(X, Y)X + v(X, Y)Y = 1$$

entonces sustituyendo la X y la Y por 0 se obtendría un absurdo.

En consecuencia, las propiedades 1, 2 y 3 enunciadas en el § 103, equivalentes en un anillo principal (capítulo 5, ej. 105) no lo son en general en el anillo de los polinomios con varias indeterminadas, lo mismo ocurre con las propiedades 4, 5 y 6 enunciadas en el mismo párrafo (ver ej. 1 más abajo).

El estudio de los ideales de $K[X_1, ..., X_m]$ ($m \geq 2$) es una parte importante del álgebra que excede infinitamente el margen de esta obra. Nos contentaremos en dar algunas propiedades elementales que nos serán útiles en lo que sigue y, por consiguiente, limitándonos a $K[X, Y]$ o $K[X, Y, Z]$ para simplificar la escritura: algunas de ellas se prolongan por sí mismas a $K[X_1, ..., X_m]$.

Si f es divisible por g en $K[X, Y]$, es evidente que f es divisible por g en $K[X][Y]$, o en $K[Y][X]$.

En consecuencia, para ver si f es divisible por g, se podrá efectuar la división euclídea de f por g en $K[X][Y]$: si los coeficientes del cociente son elementos de $K[X]$ (es decir, son polinomios en X y no fracciones racionales en X, ver capítulo 12) y si el resto es nulo, f será divisible por g (ver ej. 2, 3 y 4 más abajo).

Naturalmente esto se extiende a $K[X_1, ..., X_m] = K[X_1, ..., X_{m-1}][X_m]$.

En particular la división euclídea de $f(X, Y)$ por $X - Y$ en $K[Y][X]$ es siempre posible: el mecanismo es formalmente idéntico a la división de $f(X)$ por $X - a$ (ver § 189); se tendrá, pues,

$$f(X, Y) = (X - Y)\, q(X, Y) + r(Y)$$

siendo q y r los polinomios con coeficientes en K.

La divisibilidad de f por $X - Y$ equivale a $r = 0$, es decir, $f(X, X) = 0$, de donde:

TEOREMA 18. — *$f(X, Y)$ de $K[X, Y]$ es divisible por $X - Y$ si y sólo si $f(X, X) = 0$.*

Supongamos que f de $K[X, Y, Z]$ sea divisible por $(Y - Z)$, $(Z - X)$, $(X - Y)$, se tendrá

$$f(X, Y, Z) = (Y - Z)\, g(X, Y, Z), \qquad g \in K[X, Y, Z]$$

ahora bien, f es divisible por $Z - X$, luego $f(Z, Y, Z) = 0$ lo que implica, pues $Y - Z$ es no nulo y el anillo $K[X, Y, Z]$ íntegro, $g(Z, Y, Z) = 0$, luego

$$g(X, Y, Z) = (Z - X)\, h(X, Y, Z), \qquad h \in K[X, Y, Z]$$

igualmente $f(X, X, Z) = 0$ implica $g(X, X, Z) = 0$ y $h(X, X, Z) = 0$, de donde

$$h(X, Y, Z) = (X - Y)\, q(X, Y, Z), \qquad q \in K[X, Y, Z]$$
$$f(X, Y, Z) = (Y - Z)(Z - X)(X - Y)\, q(X, Y, Z).$$

COROLARIO. — *Si el polinomio f de $K[X, Y, Z]$ es divisible por $Y - Z$, $Z - X$, $X - Y$, es divisible por $(Y - Z)(Z - X)(X - Y)$.*

Este resultado se verifica en el caso de m indeterminadas y en el caso en que los coeficientes se toman en un anillo íntegro unitario A (ver ej. 5 más abajo).

Anotemos finalmente un resultado relativo únicamente a los *polinomios homogéneos con dos indeterminadas.* Sea

$$F(X, Y) = a_0Y^n + a_1XY^{n-1} + \ldots + a_hX^hY^{n-h} + \ldots + a_nX^n$$

las aplicaciones φ y ψ definidas en el § 186, *e*), nos permiten escribir, si $a_n \neq 0$, y suponiendo K algebraicamente cerrado, para simplificar,

$$f = \psi(F) = F(X, 1) = a_n(X - \alpha_1) \ldots (X - \alpha_n)$$

siendo $\alpha_1, \ldots, \alpha_n$ las raíces de K de F(X, 1).

Ahora bien, $a_n \neq 0$ implica que φ y ψ son biyectivas y $\varphi^{-1} = \psi$, de donde

$$F(X, Y) = \varphi(f) = a_n(X - \alpha_1Y) \ldots (X - \alpha_nY).$$

Si $a_n = 0$, se tendrá $F = Y^{n-l}G$, con

$$G = a_0Y^l + \ldots + a_hX^hY^{l-h} + \ldots + a_lX^l\, a_l$$

siendo no nulo, se factorizará a G como anteriormente a F.

Si K no es algebraicamente cerrado se tendrá

$$F(X, Y) = (X - \alpha_1Y) \ldots (X - \alpha_mY)\, Q(X, Y)$$

siendo $\alpha_1, \ldots, \alpha_m$ las raíces de F(X, 1) en K y siendo Q un polinomio de K[X, Y] tal que Q(X, 1) está desprovisto de raíces en K.

EJERCICIOS

1. Siendo f un polinomio de $K[X_1, \ldots, X_m]$, el conjunto de las $x = (x_1, \ldots, x_m)$ de K^m tal que $f(x_1, \ldots, x_m) = 0$ se llama una *hipersuperficie algebraica* de K^m (se dice *curva* para $m = 2$ y *superficie* para $m = 3$).

Si el polinomio f es *reducible*, es decir, si $f = f_1f_2$ ($0 < \text{grd } f_1 < \text{grd } f$) se dice que la hipersuperficie $f = 0$ se *descompone* en dos hipersuperficies $f_1 = 0$, $f_2 = 0$.

Demostrar que, dados f y g, si existen u y v tales que $uf + vg = 1$, no solamente que f y g son irreducibles, sino las hipersuperficies $f = 0$ y $g = 0$ tienen una intersección vacía en K^m (si K es algebraicamente cerrado se demuestra que la intersección de $f_1 = 0, \ldots, f_n = 0$ es vacía si y sólo si existe $u_1, \ldots, u_n$ tales que $u_1f_1 + \ldots + u_nf_n = 1$).

2. Demostrar que en Z[X, Y, Z], $f = X^3 + Y^3 + Z^3 - 3XYZ$ es divisible por $X + Y + Z$, hallar el cociente. Demostrar que en C[X, Y, Z], f es el producto de tres polinomios homogéneos de primer grado (ver capítulo 9, ej. 236).

3. Estudiar la divisibilidad de $X^n - Y^n$, $X^n + Y^n$ por $X - Y$ o $X + Y$ en Z[X, Y].

4. Demostrar que $f = nX^{n+1} - (n + 1)X^nY + Y^{n+1}$ es divisible por $(X - Y)^2$ en Z[X, Y], determinar el cociente.

5. Demostrar que si $f \in A[X_1, \ldots, X_m]$ es divisible por $X_i - X_j$ $(1 \leq i < j \leq n)$, f es divisible por el producto de estos polinomios (A anillo unitario íntegro).

6. Se dice que f de A[X, Y] (A anillo unitario íntegro de característica nula) es *simétrico* (resp. *antisimétrico*) si $f(Y, X) = f(X, Y)$ (resp. $f(Y, X) = -f(X, Y)$).

Demostrar que todo polinomio f antisimétrico es divisible por $X - Y$, ¿qué se puede decir del cociente? Deducir de ello que todo polinomio simétrico divisible por $X - Y$ es divisible por $(X - Y)^2$, ¿qué se puede decir del cociente?

7. Factorizar los polinomios $(X + Y)^n - X^n - Y^n$ para $n = 3$, 5 o 7 en Z[X] y en C[X].

196. Espacio vectorial y álgebra $K[X_1, ..., X_m]$

Se ve fácilmente que cualesquiera que sean f y g de $K[X_1, ..., X_m]$ y λ del cuerpo conmutativo K, $f + g$, λf pertenece a $K[X_1, ..., X_m]$ y que estas dos operaciones verifican los axiomas V_1 al V_8 que definen un espacio vectorial sobre K (ver § 125). Por otra parte, $K[X_1, ..., X_n]$ tiene una estructura de anillo y finalmente $(\lambda f)g = f(\lambda g) = \lambda(fg)$; como para $K[X]$ tenemos, pues:

TEOREMA 5'. — *Si* K *es un cuerpo conmutativo,* $K[X_1, ..., X_m]$ *provisto de la adición* $(f, g) \to f + g$ *y de la multiplicación externa* $(\lambda, f) \to \lambda f$, *tiene una estructura de espacio vectorial sobre* K; *provisto, además, de la multiplicación interna* $(f, g) \to fg$, $K[X_1, ..., X_m]$ *tiene una estructura de álgebra sobre* K.

El teorema (2') del § 186 demuestra que la familia (X^iY^j) $((i, j) \in \mathbf{N}^2)$ es una base de $K[X, Y]$, se le llama la *base canónica*; igualmente lo es la $(X_1^{i_1}, ..., X_m^{i_m})$ $((i_1, ..., i_m) \in \mathbf{N}^m)$ para $K[X_1, ..., X_m]$.

Observemos que $(m > 1)$, $K[X_1, ..., X_{m-1}]$ es un anillo y no un cuerpo, luego $K[X_1, ..., X_m]$ que tiene una estructura de *espacio vectorial* sobre K tiene una estructura de *módulo sobre el anillo* $K[X_1, ..., X_{m-1}]$ (ver capítulo 7, ej. 179).

Llamando todavía, por abuso de lenguaje, *conjunto de polinomios de grado total o igual a lo sumo a n,* al conjunto de los polinomios no nulos de grado total o igual a lo sumo a n y del polinomio cero vemos inmediatamente:

COROLARIO 1. — *El conjunto de los elementos de* $K[X_1, ..., X_m]$ *de grado igual a lo sumo a n, es un subespacio vectorial de* $K[X_1, ..., X_m]$.

Igualmente designando por conjunto los polinomios homogéneos de grado total n, la reunión de los polinomios homogéneos no nulos de grado n y del polinomio nulo, se tiene:

COROLARIO 2. — *El conjunto de los polinomios homogéneos de grado total n, es un subespacio vectorial de* $K[X_1, ..., X_m]$.

EJERCICIOS

1. ¿Cuál es la dimensión de $\mathcal{H}_n$, conjunto de los polinomios homogéneos de grado total n *de* $K[X, Y]$? (Utilizar el teorema 2' del § 186 y el ej. 3).

2. Si $\mathcal{S}_n$ designa el conjunto de los polinomios de grado total igual a lo sumo n de $K[X, Y]$ demostrar que

$$\mathcal{S}_n = \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus ... \oplus \mathcal{H}_n$$

deducir la dimensión de $\mathcal{S}_n$.

3. Aplicar los resultados de los ejercicios 1 y 2 anteriores a $K[X_1, ..., X_m]$.

4. Demostrar que $K[X]$ es un subespacio vectorial y una subálgebra del espacio vectorial o del álgebra $K[X, Y]$.

Designando, respectivamente, por $\mathcal{S}_n(X)$ y $\mathcal{S}_n(Y)$ el conjunto de los polinomios de $K[X]$ y de $K[Y]$ de grado a lo sumo igual a n, ¿cuáles son las dimensiones de $\mathcal{S}_n(X) + \mathcal{S}_n(Y)$ y de $\mathcal{S}_n(X) \cap \mathcal{S}_n(Y)$?

197. Derivación en $K[X_1, ..., X_m]$. Aplicaciones

a) Polinomios derivadas parciales

Siendo f un elemento de $K[X, Y]$ y X_1 una nueva indeterminada, distinta de X y de Y, consideremos el polinomio $f(X + X_1, Y)$ de $K[X, Y, X_1]$ podremos escribirlo (ver § 186)

$$f(X + X_1, Y) = f_0(X, Y) + f_1(X, Y)X_1 + ... + f_n(X, Y)(X_1)^n.$$

Se ve que $f_0(X, Y) = f(X, Y)$ y que el polinomio $f_1(X, Y)$ está bien determinado por el hecho de que $f(X + X_1, Y) - f(X, Y) - f_1(X, Y)X_1$ es divisible por $(X_1)^2$ en $K[X, Y, X_1]$, lo que escribimos

(1) $$f(X + X_1, Y) \equiv f(X, Y) + X_1 f_1(X, Y) \qquad (\text{mod } (X_1)^2)$$

tenemos, pues, la definición, análoga a la dada en el § 188 para $K[X]$.

DEFINICIÓN 5'. — *El polinomio* $f_1(X, Y)$, *coeficiente de* X_1 *en el polinomio* $f(X + X_1, Y)$, *considerado como elemento de* $K[X, Y, X_1]$, *se llama* polinomio derivado parcial *relativamente a* X *y se representa* f'_X; *la aplicación de* $K[X, Y]$ *en* $K[X, Y]$ *definida por* $f \to f'_X$ *se llama* derivación parcial *con relación a* X *y se representa* D_X.

Se tiene, pues, $D_X(f) = f'_X$. Se escribe también $f'_X = \dfrac{\partial f}{\partial X}$.

Se definiría igualmente la derivación parcial relativa a Y en $K[X, Y]$ y de una manera más general la derivación parcial relativamente a X_i representada D_i en $K[X_1, ..., X_m]$; siendo $D_i(f) = f'_{X_i}$ el coeficiente de Y_i en

$$f(X_1, ..., X_{i-1}, X_i + Y_i, X_{i+1}, ..., X_m)$$

considerado como elemento de $K[X_1, ..., X_i, ..., X_m, Y_i]$. Como para la derivación D en $K[X]$ tendremos:

TEOREMA 7. — *Cualesquiera que sean* f *y* g *de* $K[X_1, ..., X_m]$ *y* λ *de* K

(2) $$D_i(f + g) = D_i(f) + D_i(g), \qquad D_i(\lambda f) = \lambda D_i(f)$$

(3) $$D_i(fg) = D_i(f)g + fD_i(g).$$

Las derivaciones parciales D_i $(1 \leq i \leq m)$ *son endomorfismos de* $K[X_1, ..., X_m]$ *espacio vectorial sobre* K.

Las fórmulas (2) y (3) demuestran que D_i es una *derivación* en el álgebra $K[X_1, ..., X_m]$ sobre K, tal como la hemos definido en el ejercicio 168 del capítulo 7.

Si $i_1, i_2, ..., i_p$ son p enteros distintos o no de $[1, m]$, se podrá definir

$$D_{i_1} \circ D_{i_2} ... \circ D_{i_p},$$

llamada *derivación parcial* p-ésima. Las notaciones están muy simplificadas por el resultado siguiente:

TEOREMA 19. — *Las derivaciones* D_i *y* D_j *permutan, es decir, cualesquiera que sean* i *y* j *de* $[1, m]$

$$D_i \circ D_j = D_j \circ D_i.$$

Nos basta evidentemente verificar este resultado en $K[X, Y]$. Gracias al teorema 2′ (§ 186) y a las igualdades (2) anteriores nos basta verificar

$$(D_Y \circ D_X)(X^h Y^k) = (D_X \circ D_Y)(X^h Y^k).$$

Si h o k es nulo los dos miembros de esta igualdad son iguales, siendo nulos los dos. Si $h \geq 1$ y $k \geq 1$

$$(D_Y \circ D_X)(X^h Y^k) = D_Y(hX^{h-1}Y^k) = hkX^{h-1}Y^{k-1}$$
$$(D_X \circ D_Y)(X^h Y^k) = D_X(kX^h Y^{k-1}) = hkX^{h-1}Y^{k-1}.$$

Luego si se efectúa en $K[X, Y]$, derivaciones parciales p veces respecto a X y q veces respecto a Y, el orden de las derivaciones es indiferente, el resultado se representará $f^{(p+q)}_{X^p Y^q} = [(D_X)^p \circ (D_Y)^q](f)$, donde $(D_X)^p$ está definido por $(D_X)^1 = D_X$ y para $p > 1$, $(D_X)^p = (D_X)^{p-1} \circ D_X$. Para la generalidad de ciertos resultados se conviene que $(D_X)^0$ es la aplicación idéntica.

Así en $K[X, Y]$, f tendrá tres polinomios derivadas parciales segundas representados

$$f''_{X^2}, \quad f''_{XY}, \quad f''_{Y^2}$$

cuatro polinomios derivadas parciales de orden tres

$$f'''_{X^3}, \quad f'''_{X^2Y}, \quad f'''_{XY^2}, \quad f'''_{Y^3}$$

etcétera.

En $K[X_1, \ldots, X_m]$ el polinomio f derivada parcial p_i veces relativamente a X_i, i describiendo $[1, m]$ se representará por una de las expresiones siguientes

$$[(D_{X_1})^{p_1} \circ \ldots \circ (D_{X_m})^{p_m}]\, f = f^{(p_1 + \ldots + p_m)}_{X_1^{p_1} \ldots X_m^{p_m}} = \frac{\partial^{p_1 + \ldots + p_m}}{\partial X_1^{p_1} \ldots \partial X_m^{p_m}} f.$$

Si se pone $p = p_1 + p_2 + \ldots + p_m$ se dice que es un *polinomio derivada parcial* p-ésima, o bien de orden p.

Se ve en particular que si el *grado total* de f es n, el grado total de un polinomio derivada p-ésima de f, $p \leq n$, es a lo sumo de grado total $n - p$ y que es de grado total $n - p$ si el cuerpo K es de característica 0.

Además, si $p > n$ todo polinomio derivada p-ésima de f es nulo.

EJEMPLO

Si a, b, c son tres elementos del cuerpo conmutativo K de característica nula, calcu-lemos $f^{(p+q+r)}_{X^p Y^q Z^r}(a, b, c) = \lambda$ con $f = (X - a)^h (Y - b)^k (Z - c)^l$. Si $p + q + r \leq h + k +$ [illegible] el polinomio derivado $(p + q + r)$-ésima es nulo; si $p + q + r \leq h + k + l$, salvo un caso $(p, q, r) = (h, k, l)$, se tendrá $p < h$ o $q < h$ o $r < l$, luego el polinomio $f^{(p+q}_{X^p Y}$ [illegible] contendrá como factor una potencia estrictamente positiva de $X - a$ o de $Y - b$ o $X - c$ y λ será nulo.

Si $(p, q, r) = (h, k, l)$, $f^{(h+k+l)}_{X^hY^kZ^l}$ es igual a $h!k!l!$, luego $(p, q, r) \neq (h, k, l)$, $f^{(p+q+r)}_{X^pY^qZ^r}(a, b, c) = 0$, $f^{(h+k+l)}_{X^hY^kZ^l} = h!k!l!$

b) Derivación de los polinomios homogéneos: Fórmula de Euler

Razonemos para simplificar sobre la escritura en K[X, Y, Z], siendo K un cuerpo conmutativo de *característica nula.* Si $h(X, Y, Z)$ es un polinomio homogéneo de grado total n, se ve que h'_X, h'_Y, h'_Z son polinomios homogéneos de grado total $n - 1$ y más generalmente, para $p \leq n$, se ve que *todo polinomio derivada parcial de orden p es homogéneo de grado total* $n - p$.

Consideremos la aplicación θ de K[X, Y, Z] en sí mismo definida por

$$h \to \theta(h) = Xh'_X + Yh'_Y + Zh'_Z$$

vemos que θ es un endomorfismo del espacio vectorial K[X, Y, Z], para calcular $\theta(h)$ nos basta, pues, calcular $\theta(X^iY^jZ^k)$ con $i + j + k = n$, si $i \geq 1$, $j \geq 1$, $k \geq 1$ tendremos

$$\begin{aligned}\theta(X^iY^jZ^k) &= X(iX^{i-1}Y^jZ^k) + Y(jX^iY^{j-1}Z^k) + Z(kX^iY^jZ^{k-1}) \\ &= (i + j + k)X^iY^jZ^k = nX^iY^jZ^k.\end{aligned}$$

Se comprueba que esta fórmula es válida si uno o dos de los exponentes i, j, k son nulos. Luego para todo polinomio homogéneo de grado n se tendrá $\theta(h) = nh$; si $n > 0$ vemos que la aplicación inducida por θ sobre $\mathcal{H}_n$ espacio vectorial de los polinomios homogéneos de grado n es una *homotecia* de relación n.

Recíprocamente sea f un polinomio de K[X, Y, Z] verificando $\theta(f) = nf$.

Este polinomio puede escribirse, siendo h_i su parte homogénea de grado i (ver § 186, *b*),

$$f = \sum_i h_i$$

de donde utilizando la propiedad de f y las propiedades de las h_i

$$\theta(f) = nf = \sum_i nh_i = \sum_i \theta(h_i) = \sum_i ih_i$$

de donde, siendo única la descomposición $f = \sum_i h_i$, para todo $i \neq n$, $h_i = 0$ y $f = h_n$.

TEOREMA 20. — *Siendo* K *un cuerpo conmutativo de característica nula:*

1. *Todo polinomio de* $K[X_1, \ldots, X_m]$ *homogéneo de grado total* n *tiene como polinomios derivadas parciales de orden* p $(p \leq n)$ *polinomios homogéneos de grado total* $n - p$.
2. *Un polinomio* f *de* $K[X_1, \ldots, X_m]$ *es homogéneo de grado total* n *si y sólo si*

$$X_1f'_{X_1} + \ldots + X_mf'_{X_m} = nf \qquad (\textit{fórmula de Euler}).$$

Observemos que las aplicaciones φ y ψ definidas en el § 186, *e*) permiten obtener una "fórmula de EULER" para todo polinomio (ver ej. 2 a continuación).

c) **Fórmula de Taylor**

Si a, b, c son elementos de K, se puede escribir $X = (X - a) + a$, $Y = (Y - b) + b$, $Z = (Z - c) + c$; luego todo polinomio f de K[X, Y, Z] puede escribirse

$$f = \sum_h \sum_k \sum_l \lambda_{hkl}(X - a)^h (Y - b)^k (Z - c)^l$$

donde los elementos de K, λ_{hkl}, son nulos, salvo un número finito de ellos. Calculemos $f^{(h+k+l)}_{X^hY^kZ^l}(a, b, c)$ para los dos miembros (ver anteriormente *a*) ejemplo, tendremos

$$f^{(h+k+l)}_{X^hY^kZ^l}(a, b, c) = h!k!l!\lambda_{hkl}$$

de donde suponiendo K de característica nula:

TEOREMA 8'. — *Siendo* K *un cuerpo de característica nula, si* f *es un elemento de* K[X, Y, Z] *y* a, b, c *tres elementos cualesquiera de* K *tenemos*

$$f = \sum_h \sum_k \sum_l \frac{1}{h!\,k!\,l!} \quad f^{(h+k+l)}_{X^hY^kZ^l}(a, b, c)\,(X - a)^h (Y - b)^k (Z - c)^l$$

(*fórmula de Taylor para los polinomios de* K[X, Y, Z]).

El lector la generalizará fácilmente a los polinomios de $K[X_1, \ldots, X_m]$ e igualmente a los de $A[X_1, \ldots, X_m]$, siendo A un anillo unitario, conmutativo de característica nula.

EJERCICIOS

1. Siendo f un elemento de $K[X_1, \ldots, X_n]$ y $g_1, \ldots, g_n$ n elementos de $K[Y_1, \ldots, Y_p]$ se escribe

$$g(Y_1, \ldots, Y_p) = f[g_1(Y, \ldots, Y_p), \ldots, g_n(Y_1, \ldots, Y_p)]$$

demostrar que

$$\frac{\partial g}{\partial Y_i} = \sum_{j=1}^{n} \frac{\partial f}{\partial X_j}(g_1, \ldots, g_p)\frac{\partial g_j}{\partial X_i}.$$

2. Se tiene (ver § 186, *e*) para f de $K[X_1, \ldots, X_m]$

$$\varphi(f) = F(X_1, \ldots, X_m, T), \qquad f'_T = F'_T(X_1, \ldots, X_m, 1)$$

demostrar que $(i = 1, 2, \ldots, m)$

$$f'_{X_i} = F'_{X_i}(X_1, \ldots, X_m, 1)$$

deducir una «fórmula de EULER» para los polinomios no homogéneos.

Siendo f un polinomio de $K[X]$, deducir que los sistemas de relaciones siguientes son equivalentes ($x \in K$) (K es de característica nula)

$$\begin{cases} f\ (x) = 0 \\ f'(x) = 0 \end{cases} \Leftrightarrow \begin{cases} f'\ (x) = 0 \\ f'_T\ (x) = 0 \end{cases}$$

$$\begin{cases} f\ (x) = 0 \\ f'(x) = 0 \\ f''(x) = 0 \end{cases} \Leftrightarrow \begin{cases} f''\ (x) = 0 \\ f''_{XT}(x) = 0 \\ f''_{T^2}\ (x) = 0 \end{cases}$$

teniendo $f''_{XT} = (f'_X)'_T$ y $f''_{T^2} = (f'_T)'_T$. Generalizar.

198. Ceros de un polinomio de $K[X_1, X_2, \ldots, X_m]$

DEFINICIÓN 6'. — *Se llama* cero *de un polinomio* $f(X_1, \ldots, X_m)$ *de* $K[X_1, \ldots, X_m]$ *todo elemento* $(x_1, \ldots, x_m)$ *de* K^m, *tal que* $f(x_1, \ldots, x_m) = 0$.

Poniendo $x = (x_1, \ldots, x_m)$ se escribe algunas veces, si no da lugar a confusión, $f(x) = f(x_1, \ldots, x_m)$.

Cuando K tiene una *infinidad de elementos*, hemos visto que, para f de $K[X]$, $f(x) = 0$, para todo x de K, implica $f = 0$. Vamos a demostrar que, si f pertenece a $K[X_1, \ldots, X_m]$, $f(X_1, \ldots, X_m) = 0$, para todo $(x_1, \ldots, x_m)$ de K^m, implica $f = 0$. La propiedad es cierta para $m = 1$, supongámosla cierta para $m - 1$ y pongamos

$$f = f_0 + f_1 X_m + \ldots + f_h (X_m)^h + \ldots + f_n (X_m)^n$$

donde $f_0, \ldots, f_h, \ldots, f_n$ son elementos de $K[X_1, \ldots, X_{m-1}]$ y n el grado de f relativamente a X_m. Si $f(x_1, \ldots, x_m) = 0$ para todo $(x_1, \ldots, x_m)$ de K^m el polinomio

$$f(x_1, \ldots, x_i, \ldots, x_{m-1}, X_m) \in K[X_m]$$

es nulo para todo valor x_m de K que sustituya a X_m, luego

$$(h = 0, \ldots, n) \qquad f_h(x_1, \ldots, x_{m-1}) = 0$$

y esto cualquiera que sea $(x_1, \ldots, x_{m-1}) \in K^{m-1}$, según la hipótesis de inducción $f_0 = \ldots = f_n = 0$, luego $f = 0$.

Sea dos polinomios f y g de $K[X_1, \ldots, X_m]$ tales que sus funciones polinómicas asociadas sean iguales, es decir,

$$\tilde{f} = \tilde{g} \Leftrightarrow [\forall (x_1, \ldots, x_m) \in K^m] f(x_1, \ldots, x_m) = g(x_1, \ldots, x_m)$$

aplicando el resultado precedente a $f - g$ obtenemos un resultado análogo al del teorema 14 (§ 192):

TEOREMA 14'. — *Siendo* K *un cuerpo que tiene infinidad de elementos, la aplicación* $f \to \tilde{f}$ *de* $K[X_1, \ldots, X_m]$ *en el conjunto de las funciones polinomias* $\mathcal{S}(K^m, K)$ *es un isomorfismo de anillos, de espacios vectoriales sobre* K, *de álgebras sobre* K.

No habrá en este caso (K infinito) ningún inconveniente en representar por la misma letra el polinomio f y la función polinómica asociada: esto será en particular el caso en **Q**, **R** o **C**.

OBSERVACIONES

1 Como en el caso de una indeterminada el resultado precedente es válido reemplazando K por un anillo íntegro, unitario, infinito.

2. Si $f(x_1, ..., x_m) = 0$ para todo elemento de K^m verificando ($i = 1, ..., p$), $g_i(x_1, ..., x_m) \neq 0$, donde $g_1, ..., g_p$ son p polinomios no nulos de $K[X_1, ..., X_m]$, se tiene aún $f = 0$, si K es infinito. Consideremos, en efecto, $h = fg_1, ..., g_p$, cualquiera que sea $(x_1, ..., x_m) \in K^m$, según la hipótesis o $f(x_1, ..., x_m) = 0$ o existe i tal que $g_i(x_1, ..., x_m) = 0$ luego para todo $(x_1, ..., x_m)$ de K^m es $h(x_1, ..., x_p) = 0$ y $h = 0$; pero al ser íntegro el anillo $K[X_1, ..., X_m]$ y $g_1 \neq 0, ..., g_p \neq 0$ se tiene efectivamente $f = 0$.

199. Polinomios simétricos

a) Diremos que un polinomio f de $K[X_1, ..., X_m]$ es *simétrico* si

$$f(X_1, ..., X_m) = f[X_{p(1)}, ..., X_{p(m)}] \tag{1}$$

siendo p una permutación cualquiera de $[1, m]$, es decir, un elemento de $\mathfrak{S}_m$.

Si la igualdad (1) se verifica solamente para la *transposición* cambiando i y j ($i \neq j$), se dirá con el polinomio es simétrico en X_i y X_j. Puesto que toda permutación es un producto de transposiciones (ver § 86), un polinomio f es simétrico si y sólo si lo es con relación a toda pareja de indeterminadas distintas.

Si un polinomio simétrico contiene el monomio $X_1^{h_1}X_2^{h_2} ... X_m^{h_m}$ con un cierto coeficiente α, contiene ciertamente todos los $m!$ monomios $X_{p(1)}^{h_1} ... X_{p(m)}^{h_m}$ con el mismo coeficiente α. Pero estos $m!$ monomios no son forzosamente distintos; por ejemplo, si $m = 3$ al monomio X^3Y^2Z corresponde $3! = 6$ monomios distintos, pero a X^2YZ corresponderá solamente 3 monomios distintos por permutación: él mismo, Y^2ZX y Z^2XY; finalmente XYZ siendo el mismo simétrico los 6 monomios deducidos por permutación de las indeterminadas son iguales. De donde el interés de la definición y de la notación siguiente:

Dado el monomio $u = \alpha X_1^{i_1}X_2^{i_2} ... X_m^{i_m}$, llamaremos *simetrizado* de este monomio u y lo designaremos $s(u)$ la suma de todos los monomios *distintos*, obtenidos a partir de u, permutando de todos los modos posibles las indeterminadas $X_1, ..., X_m$. Es evidente que este simetrizado $s(u)$ es simétrico y homogéneo y de grado total $i_1 + ... + i_m$. Ordenemos lexicográficamente $s(u)$ (ver § 186, *b*) sea

$$\alpha X_1^{h_1}X_2^{h_2} ... X_m^{h_m}$$

el monomio más alto de $s(u)$, es evidente que

$$h_1 \geq h_2 \geq ... \geq h_m$$

en efecto, si $h_2 > h_1$, en $s(u)$ se hallará el monomio $\alpha X_1^{h_2}X_2^{h_1}X_3^{h_3} \ldots X_m^{h_m}$ que es estrictamente más alto que $\alpha X_1^{h_1}X_2^{h_2} \ldots X_m^{h_m}$, la misma demostración para todas las otras igualdades. Escribiremos $s(u)$ bajo la fórmula simbólica

$$s(u) = \alpha \Sigma X_1^{h_1}X_2^{h_2} \ldots X_m^{h_m} \quad \text{con} \quad h_1 \geq h_2 \geq \ldots \geq h_m.$$

El grado parcial de $s(u)$ relativo a cada indeterminada es evidentemente el mismo; debido a nuestra convención de escritura, vemos que es h_1, le llamaremos el grado parcial de $s(u)$.

Observemos que ciertos exponentes $h_1, \ldots, h_m$ pueden ser nulos, según la convención de escritura adoptada, en este caso habrá $l < m$ tal que $h_1, \ldots, h_l$ sean no nulos y $h_{l+1}, \ldots, h_m$ nulos, se escribirá entonces

$$l < m \qquad s(u) = \alpha \Sigma X_1^{h_1} \ldots X_l^{h_l} \quad \text{con} \quad h_1 \geq h_2 \geq \ldots \geq h_l.$$

b) **Polinomios simétricos elementales**

Siendo X una $m + 1$-ésima indeterminada consideremos el polinomio

$$f = (X - X_1)(X - X_2) \ldots (X - X_m)$$

f es un polinomio simétrico de $K[X_1, \ldots, X_m][X]$ podemos escribirlo en la forma

$$f = X^m - X^{m-1}\Sigma_1 + \ldots + (-1)^h X^{m-h}\Sigma_h + \ldots + (-1)^m \Sigma_m$$

donde $\Sigma_1, \ldots, \Sigma_m$ son elementos de $K[X_1, \ldots, X_m]$; se ve fácilmente que estos polinomios son simétricos, se les llama los *polinomios simétricos elementales* de $K[X_1, \ldots, X_m]$, Σ_h es precisamente la suma de todos los productos de h de las indeterminadas $X_1, \ldots, X_m$ distintas, Σ_h es entonces el simetrizado de $X_1X_2 \ldots X_h$; por consiguiente, se escribirá

$$\begin{aligned} \Sigma_1 &= \Sigma X_1 \\ &\vdots \\ \Sigma_h &= \Sigma X_1X_2 \ldots X_h \\ &\vdots \\ \Sigma_m &= \Sigma X_1X_2 \ldots X_m \end{aligned}$$

está claro que Σ_h tiene tantos términos como partes de h elementos existen en $[1, m]$, es decir, C_m^h (ver § 84), así Σ_1 tiene m términos y Σ_m uno sólo.

c) Nos proponemos demostrar el teorema siguiente:

TEOREMA 21. — *Siendo f un polinomio simétrico perteneciente a* $K[X_1, X_2, \ldots, X_m]$ *existe un polinomio único g de* $K[X_1, \ldots, X_m]$ *tal que*

$$f(X_1, \ldots, X_m) = g(\Sigma_1, \ldots, \Sigma_m),$$

siendo $\Sigma_1, \ldots, \Sigma_m$ los polinomios simétricos elementales, de $K[X_1, \ldots, X_m]$.

Existencia de un polinomio $g(\Sigma_1, \ldots, \Sigma_m)$

Todo polinomio f es suma, de una manera única, de polinomios homogéneos (ver § 186, *b*); por otra parte, toda permutación de las indeterminadas dejando invariante el grado total de un monomio, es fácil ver que cada parte homogénea de un polinomio simétrico es un polinomio simétrico: basta, pues, probar la existencia de g para un *polinomio simétrico homogéneo.*

Sea, pues, f *homogéneo, simétrico de grado total* n. Ordenemos sus monomios en el orden lexicográfico y sea

$$u_1 = \alpha X_1^{h_1} \ldots X_m^{h_m}$$

el monomio de más alto grado de f (ver § 186, *b*) como lo hemos visto anteriormente

$$h_1 \geq h_2 \geq \ldots \geq h_m.$$

Consideremos entonces el polinomio v_1 definido como sigue

$$v_1(X_1, \ldots, X_m) = \alpha(\Sigma_1)^{h_1-h_2} \ldots (\Sigma_i)^{h_i-h_{i-1}} \ldots (\Sigma_{m-1})^{h_{m-1}-h_m} (\Sigma_m)^{h_m}$$

donde $\Sigma_1, \ldots, \Sigma_m$ son los polinomios simétricos elementales de $X_1, \ldots, X_m$.

El polinomio v_1 es simétrico y homogéneo de grado total

$$h_1 - h_2 + 2(h_2 - h_3) + \ldots + (m-1)(h_{m-1} - h_m) + mh_m$$
$$= h_1 + h_2 + \ldots + h_m = \text{grd } u_1$$

busquemos el monomio el más alto de $v_1(X_1, \ldots, X_m)$; según una observación hecha en el § 186, *b*), este monomio es el producto de los monomios de mayor grado de cada uno de los factores; ahora bien, para $(\Sigma_i)^{k_i} = (\Sigma X_1, X_2, \ldots, X_i)^{k_i}$, ese monomio de mayor grado es $(X_1, X_2, \ldots, X_i)^{k_i}$, luego el monomio de mayor grado de $v_1(X_1, \ldots, X_m)$ es

$$\alpha(X_1)^{h_1-h_2}(X_1X_2)^{h_2-h_3} \ldots (X_1X_2 \ldots X_{m-1})^{h_{m-1}-h_m}(X_1 \ldots X_m)^{h_m} = \alpha X_1^{h_1} X_2^{h_2} \ldots X^{h_m} = u_1.$$

Luego si consideramos $f_1 = f - v_1$, el grado máximo del monomio de mayor grado de f_1 sea u_2, que proceda de f o de v_1 es estrictamente inferior al de u_1. Empezando de nuevo la operación precedente, obtendremos $f_2 = f_1 - v_2 = f - v_1 - v_2$, siendo el grado del monomio de mayor grado de f_2 estrictamente inferior al de u_2. Obtenemos así una sucesión

$$f_i = f - v_1 \ldots - v_i$$

de *polinomios simétricos homogéneos* de grado n tales que el grado de su monomio de mayor grado decrece estrictamente.

Ahora bien, el número de los monomios de grado total n y de grado máximo estrictamente inferior al de u_1 es finito; existe, pues, un entero tal que $f_k = 0$, luego

$$f(X_1, \ldots, X_m) = v_1(\Sigma_1, \ldots, \Sigma_m) + \ldots + v_k(\Sigma_1, \ldots, \Sigma_m) = g(\Sigma_1, \ldots, \Sigma_m).$$

Propiedades del polinomio g

Busquemos el grado total p de g, vamos a ver que es igual al grado parcial de f con relación a una de las indeterminadas X_i (en efecto, siendo f simétrico, el grado parcial relativamente a una indeterminada es el mismo para todas). Observemos a este fin que

$$\begin{aligned}
\Sigma_1 &= X_1 + \Sigma'_1 \\
\Sigma_2 &= X_1 \Sigma'_1 + \Sigma'_2 \\
&\vdots \\
\Sigma_i &= X_1 \Sigma'_{i-1} + \Sigma'_i \\
&\vdots \\
\Sigma_{m-1} &= X_1 \Sigma'_{m-2} + \Sigma'_{m-1} \\
\Sigma_m &= X_1 \Sigma'_{m-1}
\end{aligned}$$

designando por $\Sigma'_1, \ldots; \Sigma'_{m-1}$ los polinomios simétricos elementales de $X_2, \ldots, X_m$. Si g_k es la parte homogénea de grado p de g tendremos

$$g_k(\Sigma_1, \ldots, \Sigma_m) = g_k(X_1 + \Sigma'_1, \ldots, X_1 \Sigma'_{i-1} + \Sigma'_i, \ldots, X_1 \Sigma'_{m-1}),$$

el término de mayor grado en X_1 de este polinomio es

$$g_k(X_1, \ldots, X_1 \Sigma'_{i-1}, \ldots, X_1 \Sigma'_{m-1}) = (X_1)^k g_k(1, \Sigma'_1, \ldots, \Sigma'_{i-1}, \ldots, \Sigma'_{m-1}),$$

luego si $g_k \neq 0$, g_k, homogéneo de grado total k en $\Sigma_1, \ldots, \Sigma_m$, es de grado parcial k relativamente a X_1 (y a X_i cualquiera que sea i de $[1, m]$) tenemos, pues, el resultado siguiente conocido bajo el nombre de *teorema del grado*: *Si el polinomio simétrico f es de grado parcial p relativamente a una de las indeterminadas X_i, el polinomio g es de grado total p.*

Por otra parte, si f es homogéneo y de grado total n se tendrá

$$f(X_1, \ldots, X_m) = g(\Sigma_1, \ldots, \Sigma_m) = \Sigma\lambda(\Sigma_1)^{i_1} \ldots (\Sigma_m)^{i_m}$$

luego cada monomio de g, considerado como monomio en $X_1, \ldots, X_m$ es de grado total

$$i_1 + 2i_2 + \ldots + mi_m = n$$

el primer miembro de la igualdad precedente se llama *peso del monomio*

$$\lambda(\Sigma_1)^{i_1} \ldots (\Sigma_m)^{i_m};$$

sabemos, pues, que el siguiente resultado conocido con el nombre de *teorema del peso*: *Si el polinomio homogéneo simétrico f es de grado total n, cada monomio de $g(\Sigma_1, \ldots, \Sigma_m)$ es de peso n, se dice que g es isobaro de peso n.*

Unicidad de $g(\Sigma_1, \ldots, \Sigma_m)$

El teorema del grado enunciado últimamente implica que

$$f(X_1, \ldots, X_m) = 0 \Leftrightarrow g(\Sigma_1, \ldots, \Sigma_m) = 0;$$

luego si

$$f(X_1, \ldots, X_m) = g_1(\Sigma_1, \ldots, \Sigma_m) = g_2(\Sigma_1, \ldots, \Sigma_m)$$

aplicando el resultado precedente a $g_1 - g_2$ se ve que $g_1 = g_2$.

EJEMPLOS

1. Sea $m = 3$ y $f = \Sigma X^2Y$; formamos

$$f_1 = f - \Sigma_1\Sigma_2 = \Sigma X^2Y - (X + Y + Z)(XY + XZ + YZ) = -XYZ$$

luego

$$\Sigma X^2Y = \Sigma_1\Sigma_2 - \Sigma_3.$$

2. Calcular g para f de $K[X_1, ..., X_m]$

$$f = \Sigma(X_1 - X_2)^2 = g(\Sigma_1, ..., \Sigma_m)$$

observemos que g es de grado total 2 (grado parcial de f) e isóbaro de peso 2 (pues f es homogéneo de grado total 2), para cada monomio de g se tiene

$$i_1 + ... + i_m \leq 2 \quad \text{y} \quad 1i_1 + 2i_2 + ... + mi_m = 2$$

las únicas posibilidades son $i_1 = 2$, $i_k = 0$ para $k \neq 1$ e $i_2 = 1$, $i_k = 0$ para $k \neq 2$, de donde

$$\Sigma(X_1 - X_2)^2 = \lambda(\Sigma X_1)^2 + \mu\Sigma X_1X_2$$

de donde igualando los coeficientes respectivos de X_1^2 y X_1X_2 en los dos miembros $\lambda = m - 1$, $\mu = -2$.

OBSERVACION

Naturalmente el cálculo no es siempre tan simple como en estos dos ejemplos: la dificultad radica, en cada etapa, en el cálculo explícito de $f_i = f_{i-1} - \nu_i$ (para encontrar el monomio de f_i de mayor grado); para los valores pequeños de m es a menudo más simple utilizar como intermediarios los polinomios simétricos

$$S_k = (X_1)^k + ... + (X_m)^k$$

suma de las potencias semejantes de las indeterminadas (ver ejercicios posteriores).

EJERCICIOS

1. *Fórmulas de Newton.* Se pone

$$S_h = (X_1)^h + ... + (X_m)^h$$

(1) $$f = (X - X_1) \dots (X - X_m) = X^m + ... + (-1)^k X^{m-k}\Sigma_k + ... + (-1)^m\Sigma_m = (X - X_i)g_i.$$

a) Demostrar que

$$f'_X(X, X_1, ..., X_m) = mX^{m-1} - (m-1)X^{m-2}\Sigma_1 + ... + (-1)^k(m-k)X^{m-k-1}\Sigma_k + ... + (-1)^{m-1}\Sigma_{m-1} = g_1 + ... + g_m.$$

b) Calcular g_i (utilizar § 189, división por $X - a$).

c) Demostrar que para $0 < h < m$

(2) $$S_h = S_{h-1}\Sigma_1 - S_{h-2}\Sigma_2 + ... + (-1)^{h-2}S_1\Sigma_{h-1} + (-1)^{h-1}h\Sigma_h.$$

d) Sustituyendo X_i en la relación (1) eventualmente multiplicada por $(X_i)^{h-m}$ y sumando miembro a miembro las igualdades obtenidas ($h - m$ fijo) demostrar que para $h \geq m$ se tiene

(3) $$S_h = S_{h-1}\Sigma_1 - S_{h-2}\Sigma_2 + ... + (-1)^{m-2}S_{h-m-1}\Sigma_{m-1} + (-1)^{m-1}S_{h-m}\Sigma_m$$

(se observará que $S_0 = m$, luego la fórmula (2) aplicada a $h = m$ da la fórmula (3) para $h = m$).

2. Para $m = 3$ calcular S_p $(1 \leq p \leq 6)$ en función de Σ_1, Σ_2, Σ_3 (utilizar las fórmulas de NEWTON).

3. Demostrar que

$$\Sigma X_1^{p_1} X_2^{p_2} = S_{p_1} S_{p_2} - S_{p_1+p_2} \qquad \text{si} \quad p_1 \neq p_2$$

$$\Sigma X_1^p X_2^p = \frac{1}{2}(S_p)^2 - S_{2p}$$

y, por recurrencia, que $\Sigma X_1^{p_1} \ldots X_m^{p_m}$ es un polinomio en S_1, S_2 ...

4. Para $m = 3$, escribir $\Sigma X^p Y^q Z^r$ bajo la fórmula de un polinomio en S_1, S_2 ...

5. Para $m = 3$, expresar los polinomios siguientes en la forma de polinomio en S_1, S_2 ... después en la de polinomio en Σ_1, Σ_2, Σ_3

$$\Sigma(X - Y^2), \qquad \Sigma X^3 Y^2, \qquad \Sigma X^4 YZ, \qquad (X - Y)^2 (X - Z)^2 (Y - Z)^2.$$

Ejercicios

N. B. — *Salvo mención contraria los polinomios tratados tienen sus coeficientes en un cuerpo conmutativo de característica nula.*

317. Calcular por recurrencia el producto

$$(1 + X)(1 + X^2)(1 + X^4) \dots (1 + X^{2^n}).$$

318. En K[X] se pone

$$f = X^2 + 2X + 3, \qquad g = 2X^2 + 3X + 1, \qquad h = 3X^2 + X + 2.$$

Calcular $f^3 + g^3 + k^3 - 3fgh$:
a) para $\mathbf{K} = \mathbf{Z}$, **b)** para $\mathbf{K} = \mathbf{Z}/3\mathbf{Z}$.

319. En **C**[X] se considera

$$f = aX^2 + bX + c, \qquad \varphi = \alpha X^2 + \beta X + \gamma$$

¿se puede calcular a, b, c, α, β, γ de manera que $f(\varphi) = \varphi(f)$?

320. Efectuar en **R**[X] las divisiones euclídeas de
a) $X^n \operatorname{sen} \varphi - X \operatorname{sen} n\varphi + \operatorname{sen} (n - 1)\varphi$.
b) $X^{n+1} \cos (n - 1)\varphi - X^n \cos n\varphi - X \cos \varphi + 1$.
c) $X^{2n} - 2X^n \cos \varphi + 1$
por $X^2 - 2X \cos \varphi + 1$. En cada caso se obtiene que el resto es nulo; se puede demostrar de un modo más rápido este resultado colocándose en **C**[X].

321. Siendo a y b distintos calcular el resto en la división euclídea de $f(X)$ por $(X - a)(X - b)$ en función de $f(a)$ y $f(b)$. Generalizar.

322. Siendo m un entero dado ($m > 1$) todo polinomio $f(X)$, puede escribirse de una mamanera única en la forma

$$f(X) = f_0(X^m) + Xf_1(X^m) + \dots + X^{m-1}f_{m-1}(X^m)$$

donde $f_0, \dots, f_{m-1}$ son polinomios.

323. Utilizando el ejercicio 322, hallar el resto de la división de $f(X)$ por $X^m - a$.

324. Hallar el m.c.d. de $X^n - a^n$ y $X^m - a^m$.

325. ¿En qué caso $X^n + a^n$ es divisible por $X^m + a^m$?

326. ¿En qué caso $X^{2n} + X^n + 1$ es divisible por $X^2 + X + 1$?

327. Demostrar que, cualesquiera que sean los enteros naturales p, q, r, $X^{3p} + X^{3q+1} + X^{3r+1}$ es divisible por $X^2 + X + 1$. Generalizar.

328. ¿En qué caso $(X + 1)^n - X^n - 1$ es divisible por $X^2 + X + 1$?

329. ¿En qué caso $f = X^{4n} - X^{3n} + X^{2n} - X^n + 1$ es divisible por

$$g = X^4 - X^3 + X^2 - X + 1?$$

330. ¿Cuál es el resto en la división euclídea de $(\cos \alpha + X \operatorname{sen} \alpha)^n$ por $X^2 + 1$?

331. Hallar el m.c.d. de los dos polinomios

$$X^6 - X^4 - X^2 - 2, \qquad X^3 - (1 + \sqrt{2})X^2 + X(1 + \sqrt{2}) - \sqrt{2}.$$

a) en $\mathbf{Q}[X]$, b) en $\mathbf{R}[X]$.

332. Hallar el m.c.d. de los dos polinomios (si existe)

$$3X^3 + X + 1, \qquad 3X^2 + 2X - 1.$$

a) en $\mathbf{Q}[X]$, b) en $(\mathbf{Z}/3\mathbf{Z})[X]$, c) en $\mathbf{Z}[X]$.

333. Hallar todos los polinomios f tales que $f(X) + 1$ sea divisible por $(X - 1)^4$ y que $f(X) - 1$ sea divisible por $(X + 1)^4$.

a) Utilizando la relación de BEZOUT (se demostrará que hay un solo polinomio f de grado ≤ 7).

b) Considerando el polinomio derivado f'.

334. Si n es un entero estrictamente positivo.

a) Demostrar que existe una pareja única de polinomios f y g de grado estrictamente inferior a n y verificando

$$(1 - X)^n f(X) + X^n g(X) = 1,$$

b) Demostrar que

$$f(X) = g(1 - X), \qquad g(X) = f(1 - X).$$

c) Demostrar que existe una constante a tal que

$$(1 - X)f'(X) - nf(X) = aX^{n-1}.$$

Deducir los coeficientes de f y el valor de a.

335. Siendo f un polinomio de $\mathbf{K}[X]$, se considera el polinomio g definido por ($a \in \mathbf{K}$)

$$g(X) = (X - a)[P'(X) + P'(a)] - 2[P(X) - P(a)].$$

a) ¿Cuál es el orden mínimo de a, raíz de g, cuando $\mathbf{K}$ es de característica nula?

b) La misma pregunta cuando $\mathbf{K}$ es de característica p.

336. Factorizar $X^5 - 1$ en $\mathbf{K}[X]$:

a) $\mathbf{K} = \mathbf{Q}$, b) $\mathbf{K} = \mathbf{R}$, c) $\mathbf{K} = \mathbf{C}$.

337. Factorizar $X^4 - X^2 + 1$ en $\mathbf{K}[X]$ tomando por $\mathbf{K}$ cada uno de los cuerpos o anillos siguientes:

a) $\mathbf{C}$, b) $\mathbf{R}$, c) anillo de los enteros de Gauss $\mathbf{Z}(i)$, d) $\mathbf{Z}/3\mathbf{Z}$, e) $\mathbf{Z}/5\mathbf{Z}$

338. Hallar el cociente de grado n en la división respecto a las potencias crecientes de f por g

$$f(X) = 1 - abX^2, \qquad g(X) = 1 - (a + b)X + abX^2.$$

339. Se considera los polinomios A, B, C con coeficientes reales o complejos, primos dos a dos tales que $A^2 + B^2 = C^2$.

a) Demostrar que $C + B$ y $C - B$ son cuadrados de polinomios de $\mathbf{C}[X]$ y dar en consecuencia la forma general de los polinomios A, B, C.

b) Determinar A y B cuando $C = z^2 + a^2$.

Demostrar que si a es real no nulo y si A y B tienen coeficientes reales, son de la forma.

$$A = (z^2 - a^2) \text{ sen } \alpha + 2az \cos \alpha$$
$$B = (z^2 - a^2) \cos (\alpha + k\pi) - 2az \text{ sen } (\alpha + k\pi)$$

siendo α una constante real cualquiera y k un entero cualquiera.

(École des Ponts et Chaussées)

340. Demostrar que todo polinomio de $\mathbf{C}[X]$ de grado $\leq n$ puede escribirse de una manera única

$$f(X) = a_0 + a_1X + a_2X(X-1) + \dots + a_nX(X-1)\dots(X-n+1).$$

¿Cómo escoger $a_0, \dots, a_n$ de manera que para todo entero x, $f(x)$ sea un entero?

341. Sea E el espacio vectorial $\mathbf{K}[X]$, a todo elemento P de E se hace corresponder ($a \in \mathbf{K}$)

$$f(P) = (X - a)[P'(X) + P'(a)] - 2[P(X) - P(a)].$$

a) Demostrar que f es un endomorfismo de E.

b) Hallar $f(E)$, $f^{-1}(0)$. (Se podrá utilizar la base (e_k), ($k \in \mathbf{N}$) de E definida por $e_k = (X - a)^k$).

342. E es el espacio vectorial de los polinomios con coeficientes reales de grado igual o mayor que $n (n > 0)$; a y b son dos enteros tales que $0 \leq a < b$. Se considera la aplicación definida por

$$P \mapsto f(P) = X^2P'' - (a + b - 1)XP' + abP.$$

a) Demostrar que f es un endomorfismo de E.

b) *Calcular* $f(X^h)$ (*h entero natural*).

c) Determinar la imagen y el núcleo de f; se distinguirán los tres casos

$$n < a < b, \qquad a < n < b, \qquad a < b < n.$$

d) Se supone $n = 3$ hallar todos los polinomios P tales que

$$X^2P'' - 2XP' + 2P = X^3 + \lambda X + 1.$$

Discutir siguiendo los valores del número real λ. (M.G.P.)

343. Sea n un entero estrictamente positivo y E el espacio vectorial de los polinomios con coeficientes reales de grado igual o menor que n; se designa por e_p el polinomio X^p ($0 \leq p \leq n$) y se considera la aplicación f, que al polinomio P hace corresponder $Q = f(P)$ definido por

$$Q(X) = P(X + 1) + P(X - 1) - 2P(X).$$

a) Demostrar que f es un endomorfismo de E.

b) Calcular $f(e_p)$; ¿cuál es su grado? Deducir el núcleo $f^{-1}(0)$, la imagen $f(E)$ de E y el rango de f.

c) Sea Q un polinomio de $f(E)$ demostrar que existe un polinomio único P tal que

$$f(P) = Q, \qquad P(0) = P'(0) = 0.$$ (M.G.P.)

344. E es el espacio vectorial de los polinomios en X con coeficientes reales y de grado *estrictamente inferior a n*; siendo $x_1, \dots, x_n$ n números reales distintos dos a dos se pone

$$H(X) = (X - x_1)(X - x_2)\dots(X - x_n)$$

1.º Calcular $H'(x_i)$ $(1 \leq i \leq n)$ y los n polinomios

$$(i = 1, \ldots, n) \qquad E_i(X) = \frac{1}{H'(x_i)} \frac{H(X)}{x - x_i}.$$

2.º Calcular $E_j(x_i)$; deducir de ello que los polinomios $E_1, \ldots, E_n$ son independientes, después que forman una base de E.
3.º a) Demostrar que la aplicación F_i de E en **R** definida por $P \mapsto F_i(P) = P(x_i)$ es una forma lineal $(1 \leq i \leq n)$.
b) Demostrar que $(F_1, \ldots, F_n)$ es la base del dual E de E, dual de $(E_1, \ldots, E_n)$.
4.º Si $t \mapsto f(t)$ es una función real continua sobre $[a, b]$, se considera la aplicación g de E en **R** definida por

$$P \mapsto g(P) = \int_a^b f(t)P(t)dt.$$

Demostrar que g es una forma lineal; dar sus coordenadas sobre la base $(F_1, \ldots, F_n)$ de E^* en la forma de integrales definidas. (M.G.P.)

345. En todo el problema se tratará de polinomios en x cuyos coeficientes son números racionales.
1.º Dado un polinomio $g(x)$, demostrar que existe un polinomio $f(x)$ y uno sólo que verifica las condiciones.

$$f(0) = 0 \qquad f(x) - f(x-1) = g(x).$$

2.º Se designa por $f_p(x)$ el polinomio obtenido resolviendo la pregunta planteada en el punto 1.º, cuando se toma $g(x) = x^p$ (p entero > 0; por definición el polinomio x^0 es la constante 1).
Demostrar que $f_p(x)$ es divisible por $x + 1$ para $p \geq 1$. Establecer para todo entero positivo n, la relación

$$f_p(n) = 1^p + 2^p + \ldots + n^p.$$

3.º Se designa por $f'_p(x)$ la derivada del polinomio $f_p(x)$.
Verificar que

$$f'_p(x) = p\, f_{p-1}(x) + f'_p(0) \qquad \text{para} \qquad p \geq 1.$$

4.º Sea $h_p(x)$ el polinomio que tiene por derivada $f_p(x)$ y que se anula para $x = 0$.
Verificar la relación

$$f_{p+1}(x) = (p+1)h_p(x) - [(p+1)h_p(1) - 1]x.$$

Calcular explícitamente $f_p(x)$ para $p = 1, 2, 3, 4$.
5.º Se considera los tres polinomios

$$\begin{aligned} P(x) &= 1 + x + x^2 + \ldots + x^n. \\ Q(x) &= 1 + 2x + 3x^2 + \ldots + (n+1)x^n. \\ R(x) &= 1 + 2^2x + 3^2x^2 + \ldots + (n+1)^2x^n. \end{aligned}$$

Calcular el coeficiente de x_n en el desarrollo del polinomio producto $P(x)Q(x)R(x)$. Se escribirá este polinomio en la forma más simple posible. (M.G.P.)
(Primeramente se podrá observar que los polinomios

$$e_k(x) = x^{k+1} - (x-1)^{k+1} (k = 0, \ldots, n)$$

forman una base del espacio vectorial de los polinomios de grado $\leq n$).

346. Sea **K** un cuerpo conmutativo y $\mathbf{L} \neq \mathbf{K}$ un supercuerpo conmutativo de **K**. Si A es una parte de **L** se designa por $\mathbf{K}[A]$ (resp. $\mathbf{K}(A)$) el subanillo (resp. el subcuerpo) de **L** engendrado por $\mathbf{K} \cup A$. Si $A = \{\alpha\}$ se escribirá, respectivamente, $\mathbf{K}[\alpha]$ y $\mathbf{K}(\alpha)$ en este último caso se dice que el cuerpo $\mathbf{K}(\alpha)$ se ha obtenido adjuntando (añadiendo)(*) α a **K**.

a) Recordar la definición de $\mathbf{K}[A]$ y $\mathbf{K}(A)$. ¿Qué se puede decir de $\mathbf{K}[\alpha]$ y $\mathbf{K}(\alpha)$ si $\alpha \in \mathbf{K}$?

b) Se pone $\mathbf{K}_1 = \mathbf{K}(A_1)$ y $\mathbf{K}_2 = \mathbf{K}(A_2)$ demostrar que

$$\mathbf{K}(A_1 \cup A_2) = \mathbf{K}_1(A_2) = \mathbf{K}_2(A_1),$$

se escribirá $\mathbf{K}(A_1 \cup A_2) = \mathbf{K}(A_1, A_2)$ y $\mathbf{K}(\{\alpha_1, \ldots, \alpha_n\}) = \mathbf{K}(\alpha_1, \ldots, \alpha_n)$.

c) Demostrar que para toda permutación p de $\{1, \ldots, n\}$

$$\mathbf{K}(\alpha_{p(1)}, \ldots, \alpha_{p(n)}) = \mathbf{K}(\alpha_1, \ldots, \alpha_n).$$

347. Se toman las mismas notaciones que en el *ejercicio precedente* y se considera la aplicación φ del anillo $\mathbf{K}[X]$ de los polinomios en X con coeficientes en **K** en el anillo $\mathbf{K}[\alpha]$ $(\alpha \notin \mathbf{K})$ definida por

$$p \mapsto \varphi(p) = p(\alpha).$$

a) Demostrar que φ es un homomorfismo de anillos. Se designará su núcleo por I.

b) Demostrar que I es un ideal primo de $\mathbf{K}[X]$ (es decir, que $\mathbf{K}[X]$ es íntegro, V. § 99, ej. 4).

c) Si $I = \{0\}$, φ es un isomorfismo: caracterizar $\mathbf{K}(\alpha)$ respecto a $\mathbf{K}[\alpha]$; se dice que α es *trascendente* sobre **K**.

d) Si $I \neq \{0\}$ demostrar las propiedades siguientes:

— I es maximal, $\mathbf{K}[X]/I$ es un cuerpo (V. cap. 5, ej. 109).

— $\mathbf{K}[\alpha] = \mathbf{K}(\alpha)$.

— Existe un polinomio irreducible sobre **K**, unitario único f de $\mathbf{K}[X]$ tal que $I = (f)$, además grd. $f = n > 1$.

— $\mathbf{K}(\alpha)$ tiene una estructura de espacio vectorial sobre **K** su dimensión es n.

Se dice que α es *algebraico de grado n* sobre **K**.

348*. Sea **K** un cuerpo conmutativo y f un polinomio irreducible de $\mathbf{K}[X]$. Se representa por (f) el ideal engendrado por f y se pone $\mathbf{K}_1 = \mathbf{K}[X]/(f)$.

a) Demostrar que $\mathbf{K}_1$ es un cuerpo conmutativo. ¿Qué se puede decir si grd. $f = 1$?

b) Demostrar que si grd. $f \geq 2$ existe un subcuerpo propio de $\mathbf{K}_1$ isomorfo a **K**, se identificará este subcuerpo propio de $\mathbf{K}_1$ con **K**.

c) Demostrar que si grd. $f \geq 2$ existe α_1 de $\mathbf{K}_1$ tal que $\mathbf{K}_1 = \mathbf{K}(\alpha_1)$ (V. ej. 346 para esta notación), que $f(\alpha_1) = 0$ y que f es reducible sobre $\mathbf{K}_1$.

d) Demostrar que hay un supercuerpo conmutativo Δ de **K** tal que en $\Delta[X]$, f, se descomponga en factores de primer grado.

e) Se toma $\mathbf{K} = \mathbf{Q}$ y $f = X^3 - 2$, factorizar f en factores irreducibles en $\mathbf{Q}[\sqrt[3]{2}]$ y en $\mathbf{Q}[j\sqrt[3]{2}]$. Indicar un cuerpo *minimal* Δ. ¿Puede ser **C** uno de ellos?

(Si $p \in \mathbf{K}[X]$, $\dot{p} \in \mathbf{K}_1$, observar que $\dot{p} \neq \dot{0}$ es equivalente a «p y f son extraños» y utilizar la igualdad de Bezout. Si $a \in \mathbf{K} \subset \mathbf{K}[X]$, $a = \dot{a} \in \mathbf{K}_1$, según b, resulta que α_1 es precisamente $\dot{X}$. (Para **d)** razonar por inducción).

(*) *N. del T.* — Verificando la *adjunción* de α a K.

349*. $\mathbf{K}$ y $\bar{\mathbf{K}}$ son dos cuerpos conmutativos tales que existe un isomorfismo φ de $\mathbf{K}$ sobre $\bar{\mathbf{K}}$. Se representa por f un polinomio irreducible sobre $\mathbf{K}$.

a) Demostrar que existe un isomorfismo Φ de $\mathbf{K}[X]$ sobre $\bar{\mathbf{K}}[X]$ tal que su restricción en $\mathbf{K}$ coincida con φ sobre $\mathbf{K}$. Se pondrá $\bar{p} = \Phi(p)$ para todo elemento p de $\mathbf{K}[X]$.

b) Demostrar que $\mathbf{K}[X]/(f)$ y $\bar{\mathbf{K}}[X]/(\bar{f})$ son isomorfos.

c) Comparar las descomposiciones de p en factores irreducibles sobre $\mathbf{K}$ y de $\bar{p}$ en factores irreducibles sobre $\bar{\mathbf{K}}$.

350*. Se propone demostrar que si se exige a uno de los cuerpos Δ (ej. 348 **d**) el ser minimal, Δ es entonces único salvo un isomorfismo.

Se supone $n = \text{grd. } f \geq 2$. Δ y $\bar{\Delta}$ son dos supercuerpos de $\mathbf{K}$ minimales tales que

$$f(X) = (X - \alpha_1) \ldots (X - \alpha_n) \quad \text{en} \quad \Delta[X].$$
$$\bar{f}(X) = (X - \bar{\alpha}_1) \ldots (X - \bar{\alpha}_n) \quad \text{en} \quad \bar{\Delta}[X].$$

a) Demostrar que $\Delta = \mathbf{K}(\alpha_1, \ldots, \alpha_n)$, $\bar{\Delta} = \mathbf{K}(\bar{\alpha}_1, \ldots, \bar{\alpha}_n)$.

b) Demostrar que $\mathbf{K}_1 = \mathbf{K}(\alpha_1)$ y $\bar{\mathbf{K}}_1 = \mathbf{K}(\bar{\alpha}_1)$ son isomorfos, sea φ este isomorfismo y Φ el isomorfismo asociado (de $\mathbf{K}_1[X]$ sobre $\bar{\mathbf{K}}_1[X]$, V. ej. 349). Demostrar que f se descompone en factores irreducibles sobre $\mathbf{K}_1[X]$ y $\bar{\mathbf{K}}_1[X]$, respectivamente, en las formas

$$f = (X - \alpha_1)(X - \beta_1) \ldots (X - \beta_p) f_1 \ldots f_q.$$
$$\bar{f} = (X - \bar{\alpha}_1)(X - \bar{\beta}_1) \ldots (X - \bar{\beta}_p) \bar{f}_1 \ldots \bar{f}_q.$$

$\beta_i (1 \leq i \leq p)$ perteneciendo a $\{\alpha_2, \ldots, \alpha_n\}$, $\bar{\beta}_i$ a $\{\bar{\alpha}_2, \ldots, \bar{\alpha}_n\}$ con $\bar{\beta}_i = \varphi(\beta_i)$ y y $\bar{f}_j = \Phi(f_j)$ $(1 \leq j \leq q)$.

c) Demostrar, finalmente, que Δ y $\bar{\Delta}$ son isomorfos.
(Utilizar ej. 346, **c**), ej. 348 y 349.)

351. Sea G un grupo conmutativo de orden n tal que para todo divisor d de n haya a lo sumo d elementos x de G verificando $x^d = e$ (e elemento neutro de G).
Sea, en fin, $\mathbf{K}$ un cuerpo conmutativo tal que exista un entero natural n tal que el polinomio $X^n - 1$ tenga n raíces en $\mathbf{K}$.
Se escribirá

$$n = p_1^{k_1} \ldots p_m^{k_m}$$

la descomposición en factores primos de n (se pondrá $q_i = p_i^{k_i}$ para $i = 1, \ldots, m$).

a) Demostrar que para todo i de $[1, m]$ tenemos a_i de G verificando

$$(a_i)^{q_i} = e \qquad \text{y} \qquad (a_i)^{q_i/p_i} \neq e$$

deducir que a_i es de orden q_i.

b) Demostrar que existe en G un elemento de orden n y en consecuencia que G es cíclico (considerar $a_1, a_2, \ldots, a_m$ y utilizar el ej. 80, cap. 4).

c) Demostrar que las n raíces del polinomio $X^n - 1$ describen un grupo cíclico G_n, subgrupo de $\mathbf{K}^*$ (observar que $X^d - 1$ tiene a lo sumo d raíces en $\mathbf{K}$). Deducir que p es primo $(\mathbf{Z}/p\mathbf{Z})^*$ es un grupo (multiplicativo) cíclico de orden $p - 1$ y que es por lo tanto isomorfo al grupo (aditivo) $\mathbf{Z}/(p-1)\mathbf{Z}$ (utilizar un teorema de Fermat, cf. § 97, ej. 2 y § 104, ej. 3).

352. Siendo A un anillo unitario íntegro, sea f y g dos polinomios de A[X], $g \neq 0$ se representa por λ el coeficiente dominante de g. Demostrar que existe un entero natural k y dos polinomios q y r de A[X] tales que

$$\lambda^k f = gq + r \qquad (r = 0 \quad \text{o} \quad \text{grd. } r < \text{grd. } g)$$

a esta operación se la llamará división euclídea generalizada en A[X].

353*. Siendo A un anillo unitario íntegro se llama polinomio irreducible de A[X] todo elemento extremal del anillo A[X] y polinomio *primitivo* de A[X] todo polinomio cuyos coeficientes sólo tienen por divisores comunes los elementos inversibles de A.

a) Demostrar que hay en A[X] polinomios irreducibles de grado cero: ¿cuáles son?

b) Demostrar que todo polinomio f de A[X] se escribe de una sola manera $f = \lambda f^*$ (salvo factores inversibles de A) con λ elemento de A y f^* polinomio primitivo; se llamará f^* polinomio primitivo asociado a f.

Demostrar que todo polinomio primitivo de grado > 0 es irreducible.

c) Se tienen los tres polinomios

$$f = \Sigma a_h X^h, \qquad g = \Sigma b_k X^k, \qquad fg = \Sigma c_l X^l.$$

Demostrar que si p elemento extremal de A divide todos los coeficientes a_h menos uno, sea a_m, y todos los coeficientes b_k menos uno, sea b_n, p no divide c_{m+n}. Deducir que el producto de dos polinomios primitivos es primitivo.

d) Demostrar que si g^* polinomio primitivo de A[X] divide $\lambda f (\lambda \in \text{A}, f \in \text{A}[X])$, g^* divide f.

354. Se dice que un anillo A es *factorial* si es íntegro, unitario y si posee las dos propiedades F_1 y F_2 (V. cap. 5, ej. 106):

F_1: Todo elemento no inversible de A es producto de un número finito de elementos extremales de A.

F_2: La descomposición anterior es única salvo el orden y a condición de poder reemplazar cada elemento extremal por un elemento asociado, es decir, con la condición de poder multiplicarlos por los elementos inversibles de A.

Demostrar que para que A, unitario, íntegro, sea factorial es necesario y suficiente que verifique F_1 y F_2':

F_2': Todo elemento extremo p que divide ab, divide a o b.

355*. Se propone demostrar que si A es un anillo factorial también lo son los anillos A[X] y A[$X_1, \ldots, X_m$]. Se supone entonces A *factorial.*

a) Demostrar por inducción que A[X] verifica F_1 (V. ej. anterior).

b) Siendo p un elemento irreducible de A que divide a fg ($f, g \in$ A[X]), demostrar que p divide f o g (escribir $f = \alpha f^*$, $g = \beta g^*$, V. ej. 353).

c) Sea p de A[X] irreducible de grado > 0 dividiendo fg ($f, g \in$ A[X]) sin dividir f. Sea m uno de los polinomios de grado mínimo del ideal engendrado por p y f. Utilizando la división euclídea generalizada de f por m (V. ej. 352) demostrar que existe un entero k y un polinomio q de A[X] tal que $\lambda^k f = mq$ (λ coeficiente dominante de m).

Deducir de ello que m^* polinomio primitivo asociado a m divide f (observar que existe u y v de A[X] tales que $m = up + vf$ y utilizar el ejercicio 353, **d**).

Efectuando la división generalizada de p por m, demostrar que m^* divide p. Deducir que m^* es un elemento inversible de A y que p divide g. Finalmente compru-

bese que A[X] es un anillo factorial hemos obtenido así un nuevo «teorema de *transferencia*» (V. § 184, nota 2). Por otra parte $A[X_1, \dots, X_m]$ para $m > 1$ es no principal: tenemos así un ejemplo de anillo factorial no principal (V. cap. 5, ej. 106).

356*. En U_n grupo de raíces n-ésimas de 1 en **C** se considera las m raíces *primitivas* $\alpha_1, \alpha_2, \dots, \alpha_m$ (V. § 120, **c**) y se llama polinomio ciclotómico de índice n ($n \geq 1$) el polinomio de $\mathbf{C}[X]$ definido por

$$n \geq 1 \qquad \Phi_n(X) = (X - \alpha_1) \dots (X - \alpha_m), \qquad \Phi_1(X) = X - 1.$$

Se propone demostrar que $\Phi_n \in \mathbf{Z}[X]$.

a) Demostrar que $m = \varphi(n)$ (V. cap. 4, ej. 74 y cap. 5, ej. 104).

b) Demostrar que si p es primo

$$\Phi_p(X) = 1 + X + \dots + X^{p-1}.$$

c) Utilizando la descomposición de $X^n - 1$ en factores irreducibles en $\mathbf{C}[X]$ demostrar que

$$X^n - 1 = \prod_{d/n} \Phi_d(X),$$

donde el producto se extiende a todos los divisores d de n, comprendidos 1 y n (utilizar cap. 5, ej. 104, **b**).

d) Utilizando la relación anterior demostrar por inducción que Φ_n pertenece a $\mathbf{Z}[X]$.

e) Calcular $X^{12} - 1$, $X^6 - 1$ después $X^6 + 1$ en función de los polinomios Φ, deducir Φ_{12}.

357. Siendo A un anillo factorial y p un elemento extremal de A se considera el polinomio

$$f(X) = a_0 + a_1X + \dots + a_nX^n \in A[X].$$

En A se escribirá $x \equiv y(p)$ si y solamente si existe k de A tal que $x - y = kp$.

a) Demostrar que si

$$a_i \equiv 0(p) \qquad (i = 0, \dots, n-1)$$

$a_n \not\equiv 0(p)$ y $a_0 \not\equiv 0(p^2)$, el polinomio f es irreducible en $\mathbf{K}[X]$, siendo **K** el cuerpo de las fracciones de A (se demostrará que es imposible que f sea el producto de dos polinomios de grado estrictamente positivo con coeficientes en A). (Criterio de EINSENSTEIN.)

b) Demostrar que si n es un entero natural primo, $\Phi_n(X) = 1 + X + \dots + X^{n-1}$ (V. ej. 356) es irreducible en $\mathbf{Z}[X]$. (Examinar

$$g(Y) = \Phi_n(Y+1) = \frac{(Y+1)^n - 1}{Y}$$

y demostrar que se le puede aplicar el criterio precedente.)

358. Siendo **K** un cuerpo de característica distinta de 2, se tienen dos polinomios f y g de $\mathbf{K}[X, Y, Z]$ tales que

$$f(-X, Y, Z) = f(X, Y, Z), \qquad g(-X, Y, Z) = -g(X, Y, Z)$$

demostrar que existen dos polinomios f_1 y g_1 de $\mathbf{K}[X, Y, Z]$ tales que

$$f(X, Y, Z) = f_1(X^2, Y, Z), \qquad g(X, Y, Z) = Xg_1(X^2, Y, Z).$$

359. Siendo a y c dos constantes reales estrictamente positivas distintas y x e y dos variables reales se pone

$$r = \sqrt{(x-c)^2 + y^2}, \qquad r' = \sqrt{(x+c)^2 + y^2}.$$

Calcular $\quad (r + r' + 2a)(r + r' - 2a)(r - r' + 2a)(r - r' - 2a).$

Deducir las ecuaciones cartesianas de la elipse y de la hipérbola (utilizar ej. 358).

360. Calcular en $\mathbf{K}[X, Y, Z]$

$$(X + Y + Z)^3 + (X - Y - Z)^3 + (Y - Z - X)^3 + (Z - X - Y)^3.$$

361. Sea $m + 1$ polinomios $f_1, \ldots, f_m, g$ de $\mathbf{K}[X]$, $g \neq 0$ y un polinomio f de $\mathbf{K}[X_1, \ldots, X_m]$. Se efectúa las m divisiones euclídeas de f_i por g, sean $r_i (1 \leq i \leq m)$ los restos, demostrar que

$$f(f_1, \ldots, f_m) \equiv f(r_1, \ldots, r_m) \qquad (\text{mod } g).$$

362. Sea f un elemento de $\mathbf{K}[X_1, \ldots, X_m]$ tal que

$$(i = 1, \ldots, m) \qquad \frac{\partial f}{\partial X_i} = 0$$

demostrar que:

a) Si $\mathbf{K}$ es de característica nula $f \in \mathbf{K}$.

b) Si $\mathbf{K}$ es de característica p $\quad f \in \mathbf{K}[(X_1)^p, \ldots, (X_m)^p]$.

363. Siendo $\mathbf{K}$ un cuerpo conmutativo demostrar que en el anillo $\mathbf{K}[X, Y]$ el ideal engendrado por X no es maximal, aunque X sea irreducible. (Se observará que $\mathbf{K}[X, Y]$ no es principal y se considerará el ideal engendrado por X e Y.)

364. En $\mathbf{Z}[X, Y]$ estudiar la divisibilidad de $(X + Y)^n - X^n - Y^n$ por $X^2 + XY + Y^2$.

365. En $\mathbf{Z}[X, Y]$ estudiar la divisibilidad de $(X + Y + Z)^n - X^n - Y^n - Z^n$ por $(Y + Z)(Z + X)(X + Y)$.

366. En $\mathbf{Z}[X, Y, Z]$ se considera $(n \geq 2)$

$$A_n = + X^n(Y - Z) + Y^n(Z - X) + Z^n(X - Y).$$

a) Demostrar que existe un polinomio B_n de $\mathbf{Z}[X, Y, Z]$ tal que

$$A_n = -(Y - Z)(Z - X) \quad (X - Y)B_n.$$

b) Calcular B_n para $n = 2, 3, 4$. Demostrar que

$$B_n = \Sigma X^p Y^q Z^r$$

estando la Σ extendida a los elementos (p, q, r) de N^3 tales que $p + q + r = n - 2$.

367. a) Determinar los polinomios homogéneos de grado n de $\mathbf{R}[X, Y]$ verificando $h''_{X^2} = h''_{Y^2}$. (Se escribirá $h = \sum_{m=0}^{n} a_m X^m Y^{n-m}$ y se demostrará que existe una relación de recurrencia entre a_m y a_{m-2}.)

b) Demostrar que los polinomios f de $\mathbf{R}[X, Y]$ tales que $f''_{X^2} = f''_{Y^2}$ son los polinomios

$$f(X, Y) = f_1(X + Y) + f_2(X - Y),$$

donde f_1 y f_2 son dos polinomios cualesquiera de $\mathbf{R}[X]$.

368. a) Determinar todos los polinomios de $\mathbf{C}[X, Y]$ tales que $f''_{X^2} + f''_{Y^2} = 0$. (Por un método análogo al del ejercicio 367 se demostrará que

$$f(X, Y) = f_1(X + iY) + f_2(X - iY),$$

donde f_1 y f_2 son dos polinomios cualesquiera de $\mathbf{C}[X]$.

b) Entre los polinomios hallados en el a) determinar todos los que tienen coeficientes reales.

12

FRACCIONES RACIONALES

I. Fracciones racionales y funciones racionales

200. Fracciones racionales

Si K es un cuerpo conmutativo, K[X] y K[X_1, ..., X_m] son anillos unitarios íntegros; los resultados obtenidos en el § 107 nos permiten enunciar:

a) DEFINICIÓN. — *Siendo* K *un cuerpo conmutativo se llama* fracción racional *de una indeterminada* X (*resp. de m indeterminadas* X_1, ..., X_m) *todo elemento del cuerpo de fracciones del anillo íntegro* K[X] (*resp.* K[X_1, ..., X_m]). *Estos dos cuerpos se representan, respectivamente, por* K(X), K(X_1, ..., X_m).

Recordemos seguidamente que es por abuso de lenguaje el que a un elemento f de K(X) o K(X_1, ..., X_m) se le llama *fracción*: de hecho, f es una *clase de equivalencia* de la que ciertas fracciones u/v (u y v polinomios, $v \neq 0$) son *representantes*; si u_1/v_1 es otro representante de f (u_1, v_1 polinomios, $v_1 \neq 0$), tendremos

$$\frac{u}{v} = \frac{u_1}{v_1} \Leftrightarrow uv_1 = u_1v.$$

Es, pues, por abuso de notación que escribimos $f = u/v$.

Los coeficientes de los polinomios u, v se llaman también *coeficientes* de f.

Si u y v tienen sólo como divisores comunes los elementos de K^*, es decir, si son extraños, diremos que la fracción u/v es *irreducible*: por el contrario, decir que f, elemento de K(X) o de K(X_1, ..., X_m), es irreducible no tiene ningún sentido.

En K(X) podemos presentar un *representante privilegiado único* de f; sea, en efecto, dos representantes u/v y u_1/v_1 *irreducibles* de f, con v y v_1 *unitarios*; la relación

$$uv_1 = u_1v$$

implica, siendo u y v extraños, así como u_1 y v_1, que v y v_1 se dividen mutuamente; existe λ de K* tal que $v_1 = \lambda v$ y $\lambda = 1$, pues v y v_1 son unitarios, luego $u = u_1$, $v = v_1$.

Todo elemento f de K(X) *admite como representante una fracción irreducible con denominador unitario y esto de una manera única.*

b) Estructuras algebraicas definidas sobre K(X) y K(X_1, ..., X_m)

Por definición K(X) y K(X_1, ..., X_m) para la adición y la multiplicación, poseen una *estructura de cuerpo.*

Por otra parte, todo polinomio u puede considerarse como el representante de un elemento de K(X) o de K(X_1, ..., X_m), basta poner $u = u/1$.

Hemos visto igualmente que K se puede considerar como una parte de K[X] y K[X_1, ..., X_m], luego

$$K \subset K[X] \subset K(X), \qquad K \subset K[X_1, \ldots, X_m] \subset K(X_1, \ldots, X_m)$$

siendo λ un elemento de K y f una fracción racional podemos definir λf, se ve fácilmente que para la adición y la multiplicación por un elemento de K, K(X) y K(X_1, ..., X_m) poseen una *estructura de espacio vectorial sobre* K y que estos dos conjuntos dotados de la suma, de la multiplicación y de la multiplicación por un elemento de K poseen una *estructura de álgebra sobre* K.

Supongamos que A sea un *anillo unitario íntegro*: lo mismo ocurre con A[X]; designemos, respectivamente, por K y A(X) sus cuerpos de fracciones. A(X) contiene K, descrito por los elementos de representante λ/μ ($\lambda \in A$, $\mu \in A^*$); luego el cuerpo A(X) contiene el cuerpo K(X), luego el anillo K[X]; por lo tanto, también el anillo A[X], pues K contiene A, y finalmente

$$A(X) \supset K(X) \supset A[X].$$

Ahora bien, A(X) cuerpo de fracciones de A[X] es el menor cuerpo conteniendo al anillo A[X] (ver § 107), como K(X) es un cuerpo la relación precedente demuestra que

$$A(X) = K(X).$$

Se vería igualmente que (A anillo íntegro, K su cuerpo de fracciones)

$$A(X_1, \ldots, X_m) = K(X_1, \ldots, X_m).$$

201. Funciones racionales

a) Funciones racionales de una variable

Sea x un elemento de K tal que, dado f de K(X), existe un representante de f sea u_1/v_1 verificando $v_1(x) \neq 0$: se dirá que *x es sustituible en f*; si u_2/v_2 es otro representante de f verificando igualmente $v_2(x) \neq 0$ se tendrá

$$\left(\frac{u_1}{v_1} = \frac{u_2}{v_2},\ v_1(x)v_2(x) \neq 0 \right) \Rightarrow \frac{u_1(x)}{v_1(x)} = \frac{u_2(x)}{v_2(x)}$$

este valor común $u_1(x)/v_1(x)$ será llamado *valor de f para x, sustituible en f, se le representa $f(x)$.*

Se observará que esta definición, "*x sustituible en f*", no implica el que para *todo* representante u/v de f, se tenga $v(x) \neq 0$; por ejemplo, 1 es sustituible en $f = u/v$ con $u = X^3 - 1$ y $v = X^2 - 1$, pues

$$f = \frac{X^2 + X + 1}{X + 1} \Rightarrow f(1) = \frac{3}{2}$$

bien que para u/v, $v(1) = 0$.

Si u_1/v_1 es un representante irreducible de f, u_1 y v_1 no tienen raíces comunes (si no u_1 y v_1 serían divisibles por $X - a$). Si u_2/v_2 es otro representante irreducible se tiene

$$u_2 = \lambda u_1, \qquad v_2 = \lambda v_1, \qquad (\lambda \in K^*).$$

En consecuencia, toda raíz de orden h de u_1 es raíz de orden h de u_2, se dirá que es una *raíz de orden h* de f y toda raíz de orden k de v_1 es raíz de orden k de v_2, se dirá que es un *polo de orden k* de f.

Dados dos elementos f y g de $K(X)$ designemos por S_f y S_g el conjunto de los valores de K sustituibles, respectivamente, en f y en g se tendrá

(1) $\quad x \in S_f \qquad (\lambda f)(x) = \lambda f(x) \qquad (\lambda \in K)$

(2) $\quad x \in S_f \cap S_g \, (f + g)(x) = f(x) + g(x), \qquad (fg)(x) = f(x)g(x).$

Esto se aplica evidentemente a todo supercuerpo conmutativo L de K, en particular se puede tomar $L = K(X)$; ahora bien, la indeterminada X es sustituible en toda fracción $f = u/v$, pues $v \neq 0$ se escribe también $v(X) \neq 0$; luego todo elemento de $K(X)$ puede escribirse

$$\frac{u(X)}{v(X)} = f(X).$$

De un modo más general si

$$f = \frac{u}{v} = \frac{a_0 + a_1X + \ldots + a_mX^m}{b_0 + b_1X + \ldots + b_nX^n}$$

y si $g = p/q \in K(X)$ $(q \neq 0)$ se verificará que la condición

$$b_0q^n + b_1pq^{n-1} + \ldots + b_np^n \neq 0$$

es independiente de los representantes u/v de f y p/q de g: es la condición para que g sea sustituible en f; se tendrá entonces

$$f(g) = \frac{a_0q^m + a_1pq^{m-1} + \ldots + a_mp^m}{b_0q^n + b_1pq^{n-1} + \ldots + b_np^n}\, q^{n-m}.$$

En el cálculo es interesante utilizar en lugar de f y g formas reducidas irreducibles.

Las fórmulas (2) demuestran, además, que el conjunto de las fracciones racionales para las cuales x es sustituible es un subanillo A_x de K(X); además, si x, sustituible en $f = u/v$ ($v(x) \neq 0$), verifica $f(x) \neq 0$, entonces $u(x) \neq 0$ y $u(x)/v(x)$ tiene por inversa $v(x)/u(x)$; se ve, pues, que x es sustituible en $1/f = v/u$ y que

(3) $$(x \in S_f,\ f(x) \neq 0) \Rightarrow (1/f)(x) = 1/f(x).$$

Se ve fácilmente que si la fracción racional f describe el subanillo A_x de K(X), el valor $f(x)$ describe un cuerpo (ver ej. 2 más abajo).

A toda fracción racional f de K(X) se puede asociar una aplicación $\tilde{f}$ de S_f en K definida por

$$(\forall\, x \in S_f) \qquad \tilde{f}(x) = f(x)$$

$\tilde{f}$ se llama *función racional asociada a* f, está definida sobre S_f.

Las relaciones (1), (2) y (3) demuestran que

(1') $$(x \in S_f) \qquad (\widetilde{\lambda f}) = \lambda \tilde{f}$$

(2') $$(x \in S_f \cap S_g)\ (\widetilde{f + g}) = \tilde{f} + \tilde{g}, \qquad (\widetilde{fg}) = \tilde{f}\tilde{g}.$$

Finalmente si $f \neq 0$ y si S'_f es la parte de S_f tal que $f(x) \neq 0$

(3') $$(x \in S'_f) \qquad (\widetilde{1/f}) = 1/\tilde{f}.$$

Si K es infinito, sea u/v un representante de f tal que $v(x) \neq 0$

$$\left[(x \in S_f), \qquad f(x) = \frac{u(x)}{v(x)} = 0\right] \Rightarrow [u(x) = 0 \quad \text{y} \quad v(x) \neq 0];$$

ahora bien, v tiene un número finito de raíces, luego cualquiera que sea el grado n de u, habrá en S_f una infinidad de valores de x, luego al menos $n + 1$, tales que $u(x) = 0$, es decir, $u = 0$ y $f = 0$.

Supongamos que, dadas f y g de K(X)

$$(\forall\, x \in S_f \cap S_g) \qquad \tilde{f}(x) = \tilde{g}(x)$$

aplicando el resultado precedente a $f - g$ se ve que: *Si* K *es un cuerpo infinito la aplicación definida por* $f \rightarrow \tilde{f}$ *es biyectiva.*

b) **Funciones racionales de varias variables**

Las propiedades precedentes se extienden en parte a $K(X_1, \ldots, X_m)$, basta reemplazar x por $x = (x_1, \ldots, x_m) \in K^m$. Si x es sustituible en f (es decir, si existe u y v tales que $f = u/v$ con $v(x_1, \ldots, x_m) \neq 0$) se definirá

$$f(x_1, \ldots, x_m) = \frac{u(x_1, \ldots, x_m)}{v(x_1, \ldots, x_m)}.$$

Esta definición se extiende a $(x_1, \ldots, x_m) \in L^m$, donde L es un supercuerpo conmutativo de K, por ejemplo, a $L = K(X_1, \ldots, X_m)$. Si $f = u/v$, $v \neq 0$ se escribe también $v(X_1, \ldots, X_m) \neq 0$, luego las indeterminadas $X_1, \ldots, X_m$ son sustituibles en f y se puede escribir

$$f = \frac{u}{v} = \frac{u(X_1, \ldots, X_m)}{v(X_1, \ldots, X_m)} = f(X_1, \ldots, X_m).$$

Si $(x_1, \ldots, x_m)$ de K^m no es sustituible en f, puede que $(x_1, \ldots, x_m)$ sea sustituible en $1/f$ se dirá entonces que $(x_1, x_2, \ldots, x_m)$ es un *polo* de la fracción f.

Si $(x_1, \ldots, x_m)$ no es sustituible, ni en f, ni en $1/f$, se dice que $(x_1, \ldots, x_m)$ es un *punto de indeterminación de f.*

Por ejemplo, para $f = \dfrac{X^2 + Y^2}{X - Y} \in K(X, Y)$, (1, 1) es un polo y (0, 0) es un punto de indeterminación.

Designando siempre por S_f el conjunto de los $x = (x_1, \ldots, x_m) \in K^m$ sustituibles en f las fórmulas (1), (2) y (3) permanecen válidas poniendo $f(x) = f(x_1, x_2, \ldots, x_m)$. Se definirá la función racional $\tilde{f}$ asociada a f, aplica S_f en K; se demostrará que la aplicación $f \to \tilde{f}$, verifica (1′), (2′) y (3′) y que si K es infinito esta aplicación es biyectiva (ver ej. 3 más abajo).

EJERCICIOS

1. Si $f = u/v$, $g = p/q$ ($u, v, p, q \in K[X]$, $v \neq 0$, $q \neq 0$), demostrar que g es sustituible en f si y sólo si $q^n v(p/q) \neq 0$ ($n = \text{grd } v$) (es preciso demostrar que esta condición es independiente de los representantes u/v de f y p/q de g escogidos). Extender este resultado a $f \in K(X_1, \ldots, X_m)$.

2. Demostrar que si f describe el subanillo A_x de K(X) de las fracciones racionales para las cuales x es sustituible, $f(x)$ describe un cuerpo conmutativo representado $K(x)$; si $x \in K$, $K(x) = K$, si x pertenece a L supercuerpo de K, x no pertenece a K y permutando con todo elemento de K, $K(x)$ es el subcuerpo de L engendrado por $K \cup \{x\}$.

3. Demostrar que si K es infinito, y $f \in K(X_1, \ldots, X_m)$, la aplicación $f \to \tilde{f}$ es biyectiva. Hacer el mismo razonamiento para $f \in K(X)$ y utilizar la observación 2 del § 198.

4. $f \in K(X_1, \ldots, X_m)$ es *homogéneo* si f admite un representante u/v tal que u y v son homogéneos; demostrar que f es homogéneo si y sólo si

$$f(YX_1, \ldots, YX_m) = Y^d f(X_1, \ldots, X_m)$$

($d = \text{grd } u - \text{grd } v$, independiente del representante escogido se llama grado de f; ver § 204, *a*).

5. Se dice que un elemento f de K(X) es *par* (resp. *impar*) si $f(-X) = f(X)$ (resp. $f(-X) = -f(X)$), siendo K de característica distinta de 2, demostrar que se obtiene de una sola manera $f = p + i$, siendo p par e i impar. Deducir que si f es *par* (resp. impar), existe g de K(X) tal que $f(X) = g(X^2)$ (resp. $f(X) = Xg(X^2)$).

Finalmente demostrar, siendo K de característica $\neq 2$, que si $a \neq 0$ es una raíz (resp. un polo) de f par o impar $-a$ es una raíz (resp. un polo) teniendo el mismo orden de multiplicidad que a.

6. Sea $f = u/v$ un elemento de $\mathbf{C}(X)$, si $\bar{u}$ y $\bar{v}$ representan, respectivamente, los polinomios conjugados de u y v (§ 193, b), demostrar que la fracción $\bar{u}/\bar{v}$ es independiente del representante u/v de f escogido, se pone $\bar{f} = \bar{u}/\bar{v}$. Demostrar que ($\lambda \in \mathbf{C}$)

$$\overline{(f+g)} = \bar{f} + \bar{g}, \qquad \overline{fg} = \bar{f}\bar{g}, \qquad \overline{\lambda f} = \bar{\lambda}\bar{f}, \qquad \bar{\bar{f}} = f$$

deducir que la aplicación $f \to \bar{f}$ es un automorfismo involutivo del cuerpo $\mathbf{C}(X)$.

Demostrar que si $a \in \mathbf{C}$ se tiene $\overline{f(a)} = \bar{f}(\bar{a})$. Deducir que si a, complejo no real, es una raíz (resp. un polo) de orden α de f, $\bar{a}$ es una raíz (resp. un polo) de orden α de $\bar{f}$. ¿Qué conclusiones podemos obtener si $f \in \mathbf{R}(X)$?

202. Derivación de fracciones racionales

Siendo K un cuerpo conmutativo de característica nula, nos proponemos extender a K(X) la derivación definida en el § 188, en K(X).

Sea $f = u/v$ un elemento de K(X), (grd $u = l$, grd $v = m$), consideremos, con Y una nueva indeterminada, la fracción

$$f(X+Y) = \frac{u(X+Y)}{v(X+Y)} = \frac{u_0(X) + u_1(X)Y + \ldots + u_l(X)Y^l}{v_0(X) + v_1(X)Y + \ldots + v_m(X)Y^m}$$

el numerador y el denominador de la última fracción son polinomios en $\mathbf{Y}$ con coeficientes en $K[X] \subset K(X)$, es decir, los elementos de $K(X)[Y]$; siendo $v_0(X) = v(X)$ no nulo, podemos efectuar la división, relativa al entero n, según las potencias crecientes de Y, tendremos

$$u(X+Y) = v(X+Y)\,[f_0(X) + f_1(X)Y + \ldots + f_n(X)Y^n] + Y^{n+1}s(X, Y)$$

siendo $f_0, \ldots, f_n$ elementos de K(X); por otra parte, s es un polinomio en $\mathbf{Y}$ con coeficientes en K(X), luego $s(X, Y)/v(X+Y) = g_n(X, Y)$ pertenece a K(X, Y), finalmente en la división de $u(X+Y)$ por $v(X+Y)$ las únicas divisiones efectuadas son las divisiones por $v_0(X) = v(X) \neq 0$, luego $s(X, 0)$ y $g_n(X, 0)$ son elementos de K(X), finalmente

$$(1) \qquad f(X+Y) = f_0(X) + f_1(X)Y + \ldots + f_n(X)Y^n + g_n(X, Y)Y^{n+1}$$

con

$$g_n(X, Y) \in K(X, Y) \quad \text{y} \quad f_0(X), \ldots, f_n(X), g_n(X, 0) \in K(X)$$

sustituyendo 0 por Y obtenemos $f_0(X) = f(X)$, de donde para $n = 1$

$$(2) \qquad \begin{cases} f(X+Y) - f(X) = f_1(X)Y + g_1(X, Y)Y^2 \\ g_1(X, Y) \in K(X, Y) \quad \text{y} \quad f_1(X), g_1(X, 0) \in K(X). \end{cases}$$

Hemos demostrado así la existencia de f_1 verificando (2) cuando se utiliza un representante u/v de f, si demostramos que (2) define f_1 de una manera única, este elemento f_1 de K(X) será independiente del representante u/v escogido; supongamos

$$(2') \qquad \begin{cases} f(X+Y) - f(X) = \varphi_1(X)Y + \psi_1(X, Y)Y^2 \\ \psi_1(X, Y) \in K(X, Y) \quad \text{y} \quad \varphi_1(X), \psi_1(X, 0) \in K(X) \end{cases}$$

tendremos

$$f_1(X) - \varphi_1(X) = Y[\psi_1(X, Y) - g_1(X, Y)]$$

de donde sustituyendo Y por 0: $f_1(X) = \varphi_1(X)$.

La única fracción así definida se llama fracción derivada de f, se la representa f′. La relación (2) demuestra que la aplicación $f \to f'$ de K(X) en sí mismo es una *prolongación de la derivación* D definida en el § 188 en K[X]. En K(X) escribiremos igualmente $f' = Df$; se podrá igualmente definir D^k poniendo $D^1 = D$ y para $k > 1$, $D^k = D^{k-1} \circ D$. Se pondrá $f^{(k)} = D^k f$.

OBSERVACION

Si $K = \mathbf{R}$ y si $\tilde{f}$ es la función racional asociada a f tendremos (para x, y y $x + y$ pertenecientes a S_f)

$$\tilde{f}(x + y) - f(x) = y\tilde{f}'(x) + y^2\tilde{g}_1(x, y)$$

el hecho de que $g_1(X, 0) \in K(X)$ implica que

$$\lim_{y \to 0} \frac{\tilde{f}(x + y) - \tilde{f}(x)}{y} = \tilde{f}'(x),$$

luego *la función derivada de la función $\tilde{f}$ es la función racional asociada a la fracción derivada f′.*

Mediante la relación (2), demostraremos en K(X), como en K[X] las igualdades siguientes (con notaciones evidentes)

$$(4) \qquad (f + g)' = f' + g', \qquad (\lambda f)' = \lambda f', \qquad (fg)' = f'g + fg'.$$

Luego D es un *endomorfismo del espacio vectorial* K(X), igualmente D^k. Por otra parte, estas fórmulas demuestran que D es una *derivación del álgebra* K(X), tal como la hemos definido en el ejercicio 168 (capítulo 7).

Por otra parte, si $f = u/v$, tendremos

$$u = fv \Rightarrow u' = f'v + fv' = f'v + \frac{uv'}{v}$$

de donde

$$(5) \qquad \left(\frac{u}{v}\right)' = \frac{vu' - uv'}{v^2}.$$

Finalmente si $f = u/v$, grd $u = l$, grd $v = m$, se tendrá, para todo a de K, tal que $v(a) \neq 0$, utilizando la fórmula de TAYLOR para los polinomios u y v (§ 188, teorema 8)

$$f(X) = \frac{u(X)}{v(X)} = \frac{\alpha_0 + \alpha_1(X - a) + \ldots + \alpha_l(X - a)^l}{\beta_0 + \beta_1(X - a) + \ldots + \beta_m(X - a)^m}$$

de donde efectuando la división de estos polinomios en $X - a = Y$ según las potencias crecientes relativas a n ($\beta_0 = v(a) \neq 0$)

$$f(X) = \lambda_0 + \lambda_1(X - a) + \ldots + \lambda_n(X - a)^n + g_n(X)(X - a)^{n+1}$$

como en la demostración de la fórmula (1) anterior se ve fácilmente que a no es un polo de g_n; se puede, pues, en la relación precedente, sustituir X por a, se obtiene así $\gamma_0 = f(a)$. La fórmula (5) demuestra que si a no es polo de f, no es polo de $f' \ldots f^{(k)}$. Derivemos la relación precedente k veces ($k \leq n$) la fórmula de derivación de un producto nos dice que

$$f^{(k)}(X) = k!\gamma_k + (X - a)h_k(X)$$

no siendo a polo de h_k, de donde —sustituyendo X por a— $\gamma_k = f^{(k)}(a)/k!$ Obtenemos así la fórmula (6) válida para todo entero $n > 0$, llamada *fórmula de Taylor*, de orden n, de las fracciones racionales

$$(6) \qquad f(X) = f(a) + f'(a)(X - a) + \ldots + \frac{f^{(k)}(a)}{k!}(X - a)^k + \ldots + \frac{f^{(n)}(a)}{n!}(X - a)^n + g_n(X)(X - a)^{n+1}.$$

EJERCICIOS

1. Demostrar la fórmula de derivación de u/v utilizando la división de $u(X + Y)$ según las potencias crecientes. Demostrar directamente que la fórmula obtenida es independiente del representante escogido (ejercicio de puro cálculo, la parte de teoría expuesta hace inútil esta demostración).
2. Calcular $D^{(k)}f$ para $f = (X - a)^{-n}$ ($n \in \mathbf{N}$).
3. Extender la noción de derivación al álgebra $K(X_1, \ldots, X_m)$.

203. Fracciones racionales simétricas

Diremos que un elemento f de $K(X_1, \ldots, X_m)$ es *simétrico*, si para toda permutación p de $[1, m]$

$$f(X_1, \ldots, X_m) = f(X_{p(1)}, \ldots, X_{p(m)}).$$

Sea u/v un representante de f, la relación $u = fv$ demuestra que si dos de los tres elementos f, u, v son simétricos, también lo es el tercero. Luego si f es simétrico, u (resp. v) no lo es, v (resp. u) tampoco lo es; existe, pues, $i \neq j$ tales que, siendo t la transposición que cambia i y j, se tiene

$$v(X_1, \ldots, X_m) \neq v(X_{t(1)}, \ldots, X_{t(m)})$$

cambiando si fuera preciso la numeración de las indeterminadas, podemos suponer que se trata de $i = 1$ y $j = 2$; tendremos, pues,

$$f = \frac{u(X_1, X_2, X_3, \ldots, X_m)}{v(X_1, X_2, X_3, \ldots, X_m)} = \frac{u(X_2, X_1, X_3, \ldots, X_m)}{v(X_2, X_1, X_3, \ldots, X_m)}$$
$$= \frac{u(X_1, X_2, X_3, \ldots, X_m - u(X_2, X_1, X_3, \ldots, X_m)}{v(X_1, X_2, X_3, \ldots, X_m) - v(X_2, X_1, X_3, \ldots, X_m)}$$

el numerador y el denominador de la última fracción son nulos cuando se sustituye X_1 por X_2; se tiene, en consecuencia (ver § 195, teorema 18),

$$f = \frac{u}{v} = \frac{(X_1 - X_2)u_1}{(X_1 - X_2)v_1} = \frac{u_1}{v_1}$$

luego si f es simétrico, pero u y v no lo son, la fracción u/v es reducible. Por otro lado, v y v_1 son no nulos, si $f \neq 0$, *lo que supondremos*, igualmente de u y u_1; por tanto,

$$\text{grd } u_1 = \text{grd } u - 1, \qquad \text{grd } v_1 = \text{grd } v - 1.$$

Si u_1 (resp. v_1) es simétrico, también lo es de v_1 (resp. u_1). Si u_1 y v_1 no son simétricos, volveremos a empezar sobre u_1 y v_1 la operación efectuada sobre u y v, al cabo de un número k de tales operaciones obtendremos $f = u_k/v_k$ con:

— O bien: grd $u_k > 0$, grd $v_k > 0$, u_k y v_k simétricos.
— O bien: grd $u_k = 0$.
— O bien: grd $v_k = 0$.

En los dos últimos casos, u_k (resp. v_k) es simétrico, luego v_k (resp. u_k) lo es también. Podemos, pues, enunciar:

Teorema. — *Para toda fracción racional simétrica, existe un representante cociente de dos polinomios simétricos.*

En la demostración hemos encontrado el resultado siguiente: *si u/v es un representante irreducible de una fracción f simétrica, los polinomios u y v son simétricos.*

Designando por $\Sigma_1, \ldots, \Sigma_m$ los polinomios simétricos elementales de $K[X_1, \ldots, X_m]$ (§ 199, *b*), el teorema 21 del § 199 y el teorema anterior nos permiten enunciar:

Teorema. — *Si f es una fracción simétrica, elemento de* $K(_1, \ldots, X_m)$, *existe un elemento g de* $K(X_1, \ldots, X_m)$ tal que

$$f(X_1, \ldots, X_m) = g(\Sigma_1, \ldots, \Sigma_m)$$

donde $\Sigma_1, \ldots, \Sigma_m$ son los polinomios simétricos elementales de $K[X_1, \ldots, X_m]$.

EJERCICIOS

1. Demostrar que la fracción g del teorema precedente es única.
2. Aplicar las fórmulas de Newton (§ 199, ej. 1) a $h < 0$.
3. Para $m = 3$, calcular $S_h(-4 \leq h \leq -1)$ en función de $\Sigma_1, \Sigma_2, \Sigma_3$.

II. Descomposición en elementos simples

Valiéndonos del hecho de que $P = K[X]$ y $F = K(X)$ son espacios vectoriales sobre el cuerpo conmutativo K, demostraremos varias propiedades que nos serán útiles en lo que se llama *«la descomposición en elementos simples de las fracciones racionales»*, y que, además, son interesantes por sí mismas.

204. Propiedades preliminares

a) Grado de una fracción racional. Parte entera

Sea f un elemento no nulo de K(X), si u/v y u_1/v_1 son dos representantes de f, la relación $uv_1 = u_1v$ implica

$$\text{grd } u - \text{grd } v = \text{grd } u_1 - \text{grd } v_1 = d$$

este *entero racional d,* independiente del representante escogido, se llama el *grado de f.* Siendo v no nulo, efectuemos la división euclídea de u por v

$$u = pv + r \qquad (r = 0 \quad \text{o} \quad \text{grd } r < \text{grd } v),$$

de donde

$$f = \frac{u}{v} = p + \frac{r}{v} \qquad [p \in P \quad \text{y} \quad (r = 0 \quad \text{o} \quad \text{grd } r < \text{grd } v)].$$

Designemos por G la parte de $F = K(X)$ descrita por g de *grado estrictamente negativo* y por $g = 0$, vamos a ver que G es un subespacio vectorial de F. En efecto, si $g_1 = u_1/v_1$ y $g_2 = u_2/v_2$ son elementos no nulos de G y si $g_1 + g_2 \neq 0$ tendremos

$$(\text{grd } u_1 < \text{grd } v_1 \quad \text{y} \quad \text{grd } u_2 < \text{grd } v_2) \Rightarrow [\text{grd } (u_1v_2 + u_2v_1) < \text{grd } (v_1v_2)],$$

luego

$$(g_1 \in G \qquad \text{y} \qquad g_2 \in G) \Rightarrow g_1 + g_2 \in G$$

la propiedad es evidente para $g_1 = 0$ o $g_2 = 0$ o $g_1 + g_2 = 0$; como, además, es claro que para todo λ de K

$$g \in G \Rightarrow \lambda g \in G$$

resulta que G es un subespacio vectorial de $F = K(X)$; por abuso de lenguaje, se le llama *subespacio de las fracciones racionales de grado estrictamente negativo.*

La relación obtenida más arriba por división euclídea demuestra que $F = P + G$; como se ve que el único elemento f, común a P y G es $f = 0$, resulta que (§ 131, *b*): $F = P \oplus G$, de donde:

LEMA 1. — *Para todo elemento f de* K(X), *se tiene de un modo único*

(1) $$f = p + g$$

siendo p un polinomio de K[X], *llamado* parte entera *de f y g un elemento de* K(X) *nulo o de grado estrictamente negativo.*

Si grd $f < 0$: $p = 0$, $g = f$; si $f \in K[X]$: $p = f$, $g = 0$.

b) Sea g un elemento no nulo de G (luego grd $g < 0$), uno de cuyos representantes es u/v_1v_2, con los polinomios v_1 y v_2 extraños, los enteros

$$n = \text{grd } u, \qquad n_1 = \text{grd } v_1, \qquad n_2 = \text{grd } v_2$$

verifican $n < n_1 + n_2$. Consideremos la reunión de las dos familias de polinomios

$$X^k v_1 \qquad (0 \leq k \leq n_2 - 1), \qquad X^h v_2 \qquad (0 \leq h \leq n_1 - 1).$$

Estos $n_1 + n_2$ polinomios son independientes; en efecto,

$$(\lambda_0 + \lambda_1 X + \ldots + \lambda_{n_2-1} X^{n_2-1})v_1 + (\mu_0 + \quad X + \ldots + \quad _{n_1-1}X^{n_1-1})v_2 = 0$$

se escribe con las notaciones evidentes

$$w_2 v_1 + w_1 v_2 = 0$$

v_1, primo con v_2, divide w_1, lo que implica $w_1 = 0$, pues grd $w_1 <$ grd v_1, igualmente $w_2 = 0$. Estos $n_1 + n_2$ polinomios son independientes, su grado es a lo sumo igual a $n_1 + n_2 - 1$: *forman, pues, una base del espacio vectorial de los polinomios de grado a lo sumo igual a* $n_1 + n_2 - 1$.

Ahora bien, u pertenece a este espacio, luego existe una familia única de escalares $(\alpha_0, \ldots, \alpha_{n_2-1}, \beta_0, \ldots, \beta_{n_1-1})$ tal que

$$u = \sum_{k=0}^{n_2-1} \alpha_k X^k v_1 + \sum_{h=0}^{n_1-1} \beta_h X^h v_2$$

es decir, poniendo

$$u_2 = \sum_{k=0}^{n_2-1} \alpha_k X^k, \qquad u_1 = \sum_{h=0}^{n_1-1} \beta_h X^h$$

tenemos de un modo único

$$u = u_2 v_1 + u_1 v_2 \Rightarrow \frac{u}{v_1 v_2} = \frac{u_1}{v_1} + \frac{u_2}{v_2}$$

con grd $u_i <$ grd v_i $(i = 1, 2)$, o $u_1 = 0$, o $u_2 = 0$; pero si la fracción es irreducible es claro que $u_1 u_2 \neq 0$, de donde:

LEMA 2. — *Si* ν_1 *y* ν_2 *son dos polinomios primos entre sí y* u *un polinomio no nulo tal que* grd $u <$ grd $(\nu_1\nu_2)$, *existe una única pareja* (u_1, u_2) *de polinomios tales que*

$$(2) \qquad \frac{u}{\nu_1 \nu_2} = \frac{u_1}{\nu_1} + \frac{u_2}{\nu_2} (\text{grd } u_i < \text{grd } \nu_i \; i = 1, 2, \quad \text{o} \quad u_1 = 0, \quad \text{o} \quad u_2 = 0).$$

Corolario. — *Siendo* (v_i) $(1 \leq i \leq m)$ *una familia de polinomios extraños dos a dos y* u *un polinomio no nulo tal que* $\operatorname{grd} u < \operatorname{grd} (v_1 \ldots v_m)$, *existe una familia única de polinomios* (u_i) $(1 \leq i \leq m)$ *tales que*

$$\frac{u}{v_1 v_2 \ldots v_m} = \frac{u_1}{v_1} + \ldots + \frac{u_m}{v_m} \tag{2'}$$

donde ciertos polinomios u_i *pueden ser nulos, mientras que* $u_i \neq 0$ *implica* $\operatorname{grd} u_i < \operatorname{grd} v_i$.

Este resultado ha sido demostrado para $i = 2$; sea $i > 2$, supongamos lo demostrado para $i - 1$ (hipótesis de inducción). Consideremos $u/(v_1 \ldots v_i)$, pongamos $v_0 = v_1 \ldots v_{i-1}$; v_0 y v_i son primos entre sí; existe, en consecuencia, u_0 y u_i únicos (lema 2) tales que

$$\frac{u}{v_0 v_i} = \frac{u_0}{v_0} + \frac{u_i}{v_i}$$

como $u_0 \neq 0$ implica $\operatorname{grd} u_0 < \operatorname{grd} v_0$ y $u_1 \neq 0$, $\operatorname{grd} u_i < \operatorname{grd} v_i$, la hipótesis de recurrencia es aplicable a u_0/v_0, el resultado es, pues, verdadero para i, siéndolo para 2 lo es para m cualquiera. Es claro naturalmente que si $u/(v_1 \ldots v_m)$ es irreducible, para todo i de $[1, m]$, $u_i \neq 0$ y u_i/v_i es irreducible.

c) Sean u y v dos polinomios no nulos, con v de grado $p \geq 1$ y u tal que

$$\operatorname{grd} u < \operatorname{grd} (v^n) = np$$

u pertenece, por lo tanto, al espacio vectorial E de los polinomios de grado a lo sumo igual a $np - 1$, luego la dimensión es np. Se dice que la familia de polinomios

$$(X^h v^k) \qquad (h = 0, \ldots, p-1, \qquad k = 0, \ldots, n-1)$$

es una base de E; en efecto,

$$\operatorname{grd} (X^h v^k) = h + kp$$

la familia comprende np polinomios de grados respectivos

$$0, 1, \ldots, p-1, p, \ldots, 2p-1, \ldots, [p - 1 + p(n-1)] = np - 1$$

forman, pues, una base de E (§ 187, corolario del teorema 6). Ahora bien, como u pertenece a E, existe una única familia (α_{hk}) de elementos de **K** tales que

$$u = \sum_{h=0}^{p-1} \sum_{k=0}^{n-1} \alpha_{hk} X^h v^k = \sum_{k=0}^{n-1} v^k \sum_{h=0}^{p-1} \alpha_{hk} X^h$$

pongamos

$$u_{n-k} = \sum_{h=0}^{p-1} \alpha_{hk} X^h$$

tendremos de una manera única

$$u = u_n + u_{n-1} v + \ldots + u_1 v^{n-1} \Rightarrow \frac{u}{v^n} = \frac{u_n}{v^n} + \ldots + \frac{u_1}{v}$$

donde ciertos polinomios u_k pueden ser nulos, pero $u_k \neq 0$ implica grd $(u_k) <$ grd v; de donde:

LEMA 3. — *Si u y v son dos polinomios no nulos tales que*

$$\text{grd } v \geq 1 \quad \text{y} \quad \text{grd } u < n \text{ grd } v$$

existe una familia única de polinomios (u_k) $(1 \leq k \leq n)$ *que satisfacen la relación*

$$(3) \qquad \frac{u}{v^n} = \frac{u_n}{v^n} + \dots + \frac{u_1}{v} \qquad (u_k \neq 0 \Rightarrow \text{grd } u_k < \text{grd } v).$$

OBSERVACION

En el enunciado de las propiedades precedentes, no se supone que las fracciones tratadas sean irreducibles; en la práctica habrá interés por escoger representantes irreducibles.

En el mismo orden de ideas, el lema 3 es válido tanto si v es o no irreducible; cuando apliquemos este lema a la descomposición en elementos simples supondremos que v es irreducible.

EJERCICIOS

1. Demostrar el lema 2 con la ayuda de la fórmula de BEZOUT (§ 190, *b*).

2. Demostrar el lema 3 y hallar los polinomios $u_n, \dots, u_1$ efectuando las divisiones euclídeas

$$u = vq_n + u_n, \qquad q_n = vq_{n-1} + u_{n-1}, \dots, q_2 = vq_1 + u_1.$$

(Ver también otro método de cálculo en el ejercicio 384 al final de este capítulo.)

205. Descomposición de una fracción racional en elementos simples

a) Caso general

Sea f un elemento de K(X) (K cuerpo conmutativo), tenemos de una manera *única* (§ 204, lema 1)

$$f = p + g \qquad (p \in K[X], \qquad g \in K(X), \qquad g = 0 \quad \text{o} \quad \text{grd } g < 0)$$

g admite un representante *único* irreducible con denominador unitario v, luego

$$f = p + \frac{u}{v} \qquad (u = 0 \quad \text{o} \quad \text{grd } u < \text{grd } v).$$

Existe una descomposición canónica de v en factores irreducibles unitarios (§ 191, teorema 11), luego

$$(1) \qquad f = p + \frac{u}{(v_1)^{\alpha_1} \dots (v_m)^{\alpha_m}}$$



con $v_1 \dots v_m$ irreducibles, unitarios y tales que

$$u \neq 0 \Rightarrow \text{grd } u < \sum_{h=1}^{n} \alpha_h \text{ grd } (v_h).$$

En lo que sigue supondremos $u \neq 0$; $(v_1)^{\alpha_1}, \dots, (v_m)^{\alpha_m}$ son extraños dos a dos: según el corolario del lema 2 (§ 204) existe una familia única (u_h) $(1 \leq h \leq m)$ de polinomios verificando

$$(2) \qquad f = p + \sum_{h=1}^{m} \frac{u_h}{(v_h)^{\alpha_h}} \qquad (h = 1, \dots, m, \text{ grd } (u_h) < \text{grd } (v_h)).$$

Finalmente según el lema 3 (§ 204), existe para todo h de $[1, m]$ una familia única (r_{hk}) $(1 \leq h \leq m)$ $(1 \leq k \leq \alpha_h)$ de polinomios verificando

$$(3) \qquad f = p + \sum_{h=1}^{m} \sum_{k=1}^{\alpha_h} \frac{r_{hk}}{(v_h)^k}$$

y

$$r_{hk} \neq 0 \Rightarrow \text{grd } (r_{hk}) < \text{grd } (v_h),$$

de donde finalmente:

Teorema. — *Para todo elemento f de* K(X) *en el que el denominador v del representante irreducible unitario tiene por descomposición canónica*

$$v = (v_1)^{\alpha_1} \dots (v_m)^{\alpha_m}$$

existe un polinomio único p y una familia única (r_{hk}) $(1 \leq h \leq m,\ 1 \leq k \leq \alpha_h)$ de polinomios tales que

$$(3) \qquad f = p + \sum_{h=1}^{m} \sum_{k=1}^{\alpha_h} \frac{r_{hk}}{(v_h)^k}$$

donde $r_{hk} \neq 0$ implica $\text{grd } (r_{hk}) < \text{grd } (v_h)$.

El segundo miembro de (3) *se llama descomposición de f en elementos simples.*

Observemos que si f no es un polinomio, todos los polinomios $r_{h\alpha_h}$ $(1 \leq h \leq m)$ son no nulos, sino tendríamos, según (3), para f un representante cuyo denominador contendría el factor v_h a una potencia estrictamente inferior a α_h, lo que es imposible.

b) Descomposición en C(X) o K(X), K cuerpo conmutativo algebraicamente cerrado

Si K es algebraicamente cerrado (§ 193), que es el caso de **C**, tendremos

$$v = (X - a_1)^{\alpha_1} \dots (X - a_m)^{\alpha_m}$$

$a_1, \ldots, a_m$ son los polos, dos a dos distintos de f, de órdenes respectivos $\alpha_1, \ldots, \alpha_m$.

Al ser de primer grado los polinomios $v_h = X - a_h$, los polinomios r_{hk} son elementos de K. Si escribimos

$$P\left(\frac{1}{X-a}\right) = \frac{A_1}{X-a} + \ldots + \frac{A_\alpha}{(X-a)^\alpha}$$

la parte de la descomposición relativa al polo a, de orden α, observamos que el polinomio P es de grado α, pues $A_\alpha \neq 0$, de lo contrario a sería un polo de orden estrictamente inferior a α, y que P no tiene término constante, luego $\omega(P) \geq 1$.

El coeficiente A_1 se llama el *residuo* relativo al polo a. Tendremos, pues, de una manera única (salvo el orden)

$$(4) \qquad f(X) = p(X) + P_1\left(\frac{1}{X-a_1}\right) + \ldots + P_m\left(\frac{1}{X-a_m}\right)$$

con $p \in K[X]$ y para $h = 1, \ldots, m$

$$P_h \in K[X], \qquad \text{grd}\ (P_h) = \alpha_h, \qquad \omega(P_h) \geq 1.$$

Las fracciones de la forma $A_k/(X-a)^k$ se llaman (veremos el por qué en el § 207) *elementos simples de primera especie.*

OBSERVACION

La fórmula (4) permite expresar f de $\mathbf{C}(X)$ como combinación lineal de elementos de $\mathbf{C}(X)$ de la forma X^h, $1/(X-a)^k$ ($a \in \mathbf{C}$, h, $k \in \mathbf{N}$). Esta *linearización* de f (en el sentido dado a esta palabra en el § 147, *c*) es muy útil cuando, siendo φ un *endomorfismo* del espacio vectorial $\mathbf{C}(X)$, se quiere calcular $\varphi(f)$: basta saber calcular $\varphi(X^h)$ y $\varphi(1/(X-a)^k)$.

Si $\mathbf{C}'$ es una parte de $\mathbf{C}$, el conjunto E de las *funciones racionales* definidas sobre $\mathbf{C}'$ y con valor en $\mathbf{C}$ es un espacio vectorial sobre $\mathbf{C}$. Gracias a la biyección asociando $\tilde{f}$ a f (siendo $\mathbf{C}$ infinito, ver § 201, *a*) φ, endomorfismo de $\mathbf{C}(X)$, puede ser considerado como un endomorfismo de E; ahora bien, para los operadores lineales siguientes (donde $[\alpha, \beta]$ de $\mathbf{R}$ está incluido en $\mathbf{C}'$):

1. Derivación. Derivación n-ésima.
2. Desarrollo limitado en el entorno de un punto de $[\alpha, \beta]$.
3. Desarrollo en serie en el entorno de un punto de $[\alpha, \beta]$.
4. Cálculo de la primitiva anulándose en un punto de $[\alpha, \beta]$.
5. Integración en $[\alpha, \beta]$.

Se verá en el curso de Análisis que conocemos las imágenes por φ de las funciones racionales asociadas a los «elementos simples» X^h y $1/(X-{}^k)$.

De ahí el gran interés de la descomposición en elementos simples de las fracciones racionales en $\mathbf{C}(X)$.

206. Métodos prácticos de descomposición en C(X)

Dada la *unicidad* de la descomposición dada por la fórmula (4) del párrafo precedente, f y $g = f - p$ *tienen los mismos polos y la misma descomposición* relativa a cada uno de estos polos; igualmente f, $f - p = g$ y

$$g_1 = f - p - P_1\left(\frac{1}{x - a_1}\right)$$

tienen la misma descomposición relativa a cada uno de los polos distintos del a_1: el orden de cálculo de los polinomios p, P_1, ..., P_m no tiene importancia; habiéndose calculado ya p, P_1, ..., P_h resulta algunas veces más simple, pero no siempre, operar sobre

$$g_h = f - p - P_1\left(\frac{1}{X - a_1}\right) \dots - P_h\left(\frac{1}{X - a_h}\right)$$

para calcular P_{h+1}, ..., P_m. *Por el contrario resulta casi siempre interesante calcular p por división euclídea y seguidamente operar sobre $g = f - p$.*

a) **Cálculo de $P\left(\frac{1}{X - a}\right)$ por división según las potencias crecientes**

Sea a un polo de orden α de u/v supuesto irreducible

$$\frac{u}{v} = \frac{u(X)}{(X - a)^\alpha v_1(X)}, \qquad v_1(a) \neq 0.$$

Pongamos, ordenando U y V_1 según las potencias crecientes

$$X = a + Y, \qquad u(X) = u(a + Y) = U(Y), \qquad v_1(X) = v_1(a + Y) = V_1(Y).$$

Como $V_1(0) = v_1(a) \neq 0$, podemos efectuar la división según las potencias crecientes de U por V_1 relativamente a todo entero n (§ 194), tomemos $n = \alpha - 1$

$$U(Y) = (A_\alpha + A_{\alpha-1}Y + \dots + A_1Y^{\alpha-1})V_1(Y) + Y^\alpha U_1(Y)$$

A_1, ..., A_α son dos números complejos. De donde, sustituyendo $X - a$ por Y y poniendo $U_1(Y) = U_1(X - a) = u_1(X)$,

$$u(X) = [A_1(X - a)^{\alpha-1} + \dots + A_\alpha]v_1(X) + (X - a)^\alpha u_1(X)$$
$$\frac{u(X)}{v(X)} = \frac{u(X)}{(X - a)^\alpha u_1(X)} = \frac{A_1}{X - a} + \dots + \frac{A_\alpha}{(X - a)^\alpha} + \frac{u_1(X)}{v_1(X)}$$

la fracción u_1/v_1 está desprovista del polo a, pues $v_1(a) \neq 0$, luego la parte de la descomposición relativa al polo es

$$P\left(\frac{1}{X - a}\right) = \frac{A_1}{X - a} + \dots + \frac{A_\alpha}{(X - a)^\alpha}.$$

b) Cálculo directo de A_α. Generalización

Siendo a un polo de orden α de u/v irreducible, sabemos que existen números complejos $A_1, \ldots, A_\alpha$ y una fracción u_1/v_1 desprovista del polo a tales que

$$(1) \qquad \frac{u}{v} = \frac{u(X)}{(X-a)^\alpha v_1(X)} = \frac{A_\alpha}{(X-a)^\alpha} + \ldots + \frac{A_1}{(X-a)} + \frac{u_1(X)}{v_1(X)}.$$

Multipliquemos los dos miembros de esta igualdad por $(X-a)^\alpha$, obtenemos $(\alpha \geq 1)$

$$(2) \qquad \frac{u(X)}{v_1(X)} = A_\alpha + (X-a)h_1(x)$$

siendo h_1 una fracción desprovista de polo a; sustituyendo X por a en (2) se obtiene, siendo x una variable compleja

$$A_\alpha = \frac{u(a)}{v_1(a)} = \lim_{x\to a}\left[(x-a)^\alpha \frac{u(x)}{v(x)}\right]$$

En particular si a es polo simple el residuo A nos lo da la fórmula

$$A = \frac{u(a)}{v_1(a)} = \frac{u(a)}{v'(a)}$$

pues $v(X) = (X-a)v_1(X)$ implica $v'(X) = v_1(X) + (X-a)v_1'(a)$.

Si $\alpha \geq 2$ se puede generalizar este método al cálculo de $A_{\alpha-1}$; en efecto, la relación (2) puede escribirse

$$(3) \qquad \frac{u(X)}{v_1(X)} = A_\alpha + A_{\alpha-1}(X-a) + (X-a)^2 h_2(X)$$

siendo h_2 una fracción desprovista del polo a. Derivemos la relación (3), obtenemos

$$(3') \qquad \left[\frac{u(X)}{v_1(X)}\right]' = A_{\alpha-1} + (X-a)k(x)$$

donde k es una fracción desprovista del polo a, de donde sustituyendo X por a en (3')

$$A_{\alpha-1} = \left[\frac{u(X)}{v_1(X)}\right]'_{X=a}.$$

Si $\alpha \geq 3$ se podría continuar este método para calcular $A_{\alpha-2}, \ldots$, etc.; pero entonces es fácil ver que si se aplica la fórmula de TAYLOR a la fracción $\varphi = u/v_1$ hasta el orden $\alpha - 1$, para $X = a$, se obtiene

$$\frac{u(X)}{v_1(X)} = \varphi(a) + (X-a)\varphi'(a) + \ldots + \frac{(X-a)^{\alpha-1}}{(\alpha-1)!}\varphi^{\alpha-1}(a) + (X-a)^\alpha \psi(x)$$

siendo ψ una fracción desprovista del polo a (§ 202, fórmula 6), en consecuencia

$$P\left(\frac{1}{X-a}\right)=\frac{\varphi(a)}{(X-a)^{\alpha}}+\frac{\varphi'(a)}{(X-a)^{\alpha-1}}+\dots+\frac{\varphi^{(\alpha-1)}(a)}{(\alpha-1)!}\frac{1}{X-a}.$$

Esta fórmula, que tiene el interés de presentar una cierta analogía entre el estudio de la parte relativa al polo a, y la fórmula de TAYLOR es poco utilizada en la práctica, ya que el cálculo de las derivadas sucesivas de una función racional es en general muy molesto; no obstante, en el caso de una fracción que tenga dos polos solamente puede conducir a un cálculo práctico de la descomposición (ver ej. 371 al final del capítulo).

c) **Método de los coeficientes indeterminados**

Se puede escribir *a priori*

$$(5) \qquad f=\frac{u}{v}=p+\sum_{h=1}^{m}\sum_{k=1}^{\alpha_h}\frac{A_{hk}}{(X-a_h)^k}$$

multiplicando los dos miembros por v, se obtiene una igualdad entre polinomios. Igualando los coeficientes de X^l en los dos miembros, se constatará, en cada caso, que se obtiene un sistema de CRAMER respecto a los coeficientes indeterminados: se demuestra de esta manera cada vez la existencia y la unicidad de la descomposición anterior.

Si se utiliza este método, conviene aplicarlo a $g=f-p$ y no a f.

Se puede también poner *a priori* la descomposición de $g=f-p$, si los polos son de orden poco elevado ($\alpha\leq 3$, por ejemplo); se calculará A_α e incluso $A_{\alpha-1}$ por el método indicado antes en *b*). Se obtendrá en seguida el número de relaciones necesarias para calcular los otros coeficientes indeterminados dando *valores particulares* a X (valores que no son polos), en particular 0, si no es polo.

Colocándose en $\mathbf{C}=\mathbf{C}\cup\{\infty\}$ (ver § 124) se podrá sustituir X por ∞ en la relación obtenida multiplicando los dos miembros de la fórmula (5) por X. Sea, por ejemplo,

$$f(X)=\frac{2X^4+1}{(X-1)^3(X^2+1)}=\frac{A_3}{(X-1)^3}+\frac{A_2}{(X-1)^2}+\frac{A_1}{X-1}+\frac{B}{X-i}+\frac{C}{X+i}$$

$$Xf(X)=\frac{A_3X}{(X-1)^3}+\frac{A_2X}{(X-1)^2}+\frac{A_1X}{X-1}+\frac{BX}{X-i}+\frac{CX}{X+i}$$

se tendrá

$$2=A_1+B+C$$

de una manera general (notaciones de la fórmula (5))

$$\lim_{x\to\infty} xg(x)=A_{11}+\dots+A_{m1}.$$

Se tendrá también en cuenta las consideraciones siguientes: si f es *par* (§ 201, ej. 5) la relación

$$f(-X) = p(-X) + g(-X) = f(X) = p(X) + g(X)$$

implica que p y g son pares; luego si $a \neq 0$ es polo de orden α, $-a$ será igualmente polo de orden α, la suma de las partes relativas a los polos a y $-a$ sea

$$P\left(\frac{1}{X-a}\right) + Q\left(\frac{1}{X+a}\right)$$

es invariante cuando se sustituye $-X$ por X, de donde

$$\sum_{k=1}^{\alpha}\left[\frac{A_k}{(X-a)^k} + \frac{B_k}{(X+a)^k}\right] = \sum_{k=1}^{\alpha}\left[\frac{A_k}{(-X-a)^k} + \frac{B_k}{(-X+a)^k}\right]$$

de donde como consecuencia de la unicidad de la descomposición

$$(k = 1, \ldots, m) \qquad B_k = (-1)^k A_k.$$

Si f es *impar* (§ 201, ej. 5), se verá sin esfuerzo, con las mismas notaciones, que

$$(k = 1, \ldots, m) \qquad B_k = (-1)^{k-1} A_k.$$

Finalmente si f es *real*: $\bar{f} = f$, igualmente con p y con g y si a no real es polo de orden α, $\bar{a}$ es de polo de orden α (ver § 201, ej. 6); siendo P y Q las partes relativas a $\bar{a}$ y a en la descomposición, tendremos

$$f(X) = P\left(\frac{1}{X-a}\right) + Q\left(\frac{1}{X-\bar{a}}\right) + f_2(X)$$

estando f_2 desprovisto de los polos a y $\bar{a}$; además, la igualdad

$$\bar{f}(X) = \bar{P}\left(\frac{1}{X-\bar{a}}\right) + \bar{Q}\left(\frac{1}{X-a}\right) + \bar{f}_2(X) = f(X)$$
$$= P\left(\frac{1}{X-a}\right) + Q\left(\frac{1}{X-\bar{a}}\right) + f_2(X)$$

implica, en consecuencia, de la unicidad de la descomposición

$$Q = \bar{P} \quad \text{y} \quad \bar{f}_2 = f_2$$

luego si se pone

$$P\left(\frac{1}{X-a}\right) = \sum_{k=1}^{\alpha} \frac{A_k}{(X-a)^k}, \qquad Q\left(\frac{1}{X-\bar{a}}\right) = \sum_{k=1}^{\alpha} \frac{B_k}{(X-\bar{a})^k}$$

se tendrá

$$(k = 1, \ldots, m) \qquad B_k = \overline{A_k}.$$

d) Conclusión

Interesa calcular la parte entera p (si hay una) por división euclídea. Para los polos de orden relativamente alto ($\alpha > 3$, por ejemplo), el método descrito anteriormente en *a*) es el mejor.

Para los polos de orden bajo ($\alpha \leq 3$, por ejemplo) se hará *a priori* la descomposición, se calculará inmediatamente A_α como se ha dicho en *b*), seguidamente la sustitución de valores particulares, junto a la utilización eventual de consideraciones de semejanza o de realidad, dará las relaciones que permitirán calcular los demás coeficientes.

El método de los coeficientes indeterminados aplicado sin meditar no es en general aconsejable: hay que reducir a común denominador el segundo miembro de (5), después resolver un sistema lineal con bastantes incógnitas, aunque los casos sean simples.

EJEMPLOS

1. Finalicemos el ejemplo dado anteriormente teniendo en cuenta que $\overline{f} = f$.

$$f(X) = \frac{2X^4 + 1}{(X-1)^3(X^2+1)} = \frac{A_3}{(X-1)^3} + \frac{A_2}{(X-1)^2} + \frac{A_1}{X-1} + \frac{B}{X-i} + \frac{\overline{B}}{X+i}$$

$$A_3 = \frac{2 \cdot 1 + 1}{1^2 + 1} = \frac{2}{2}, \qquad B = \frac{2i^4 + 1}{(i-1)^3 2i} = -\frac{3}{8}(1+i)$$

$X = \infty$ da (después de multiplicar los dos miembros por X)

$$2 = A_1 + B + \overline{B} \Rightarrow A_1 = \frac{11}{4}$$

finalmente $X = 0$ da

$$-1 = -\frac{3}{2} + A_2 - \frac{11}{4} \Rightarrow A_2 = \frac{3}{4}.$$

2. $$f(X) = \frac{X^6 - 2X^5 + 4X^4 - 6X^3 - X^2 + 8X + 121}{(X-1)^3(X^2+4)} \qquad \text{(M. G. P.)}$$

por división euclídea se obtiene

$$f(X) = X + 1 + g(X), \qquad g(X) = \frac{125}{(X-1)^3(X^2+4)}$$

pongamos

$$Y = X - 1, \qquad X^2 - 4 = 5 + 2Y + Y^2$$

$$125 = (5 + 2Y + Y^2)(25 - 10Y - Y^2) + Y^3(12 + Y)$$

$$g(X) = \frac{25}{(X-1)^3} - \frac{10}{(X-1)^2} - \frac{1}{X-1} + h(X), \qquad h(X) = \frac{X + 11}{X^2 + 4}$$

$$h(X) = \frac{X+11}{X^2+4} = \frac{A}{X-2i} + \frac{\overline{A}}{X+2i}, \qquad A = \frac{2i+11}{4i} = \frac{2-11i}{4}$$

$$f(X) = X + 1 + \frac{25}{(X-1)^3} - \frac{10}{(X-1)^2} - \frac{1}{X-1} + \frac{2-11i}{4}\left(\frac{1}{X-2i}\right) + \frac{2+11i}{4}\left(\frac{1}{X+2i}\right)$$

$$f(X) = \frac{1}{X^n - 1}; \qquad u(X) = 1; \qquad v(X) = \prod_{k=0}^{n-1} (X - a_k);$$

se tiene, pues,

$$\frac{1}{X^n - 1} = \prod_{k=0}^{n-1} \frac{A_k}{X - a_k}$$

con

$$A_k = \frac{u(a_k)}{v'(a_k)} = \frac{1}{n(a_k)^{n-1}} = \frac{a_k}{n};$$

ahora bien,

$$(a_k)^n = 1,$$

de donde

$$\frac{1}{X^n - 1} = \frac{1}{n} \sum_{k=0}^{n-1} \frac{a_k}{X - a_k} \qquad \left(a_k = \cos k \frac{2\pi}{n} + i \operatorname{sen} k \frac{2\pi}{n} \right).$$

207. Descomposición en R(X)

a) Unicidad de la descomposición

Sea f un elemento de $\mathbf{R}(X)$, u/v su representante irreducible con denominador unitario; tenemos en $\mathbf{R}[X]$ (ver § 193, *c*)

$$v(X) = \prod_{h=1}^{m} (X - a_h)^{\alpha_h} \prod_{k=1}^{n} [(X - b_k)^2 + (c_k)^2]^{\beta_k}$$

siendo (a_h) $(1 \leq h \leq m)$ la familia de los *polos reales* de f, de órdenes respectivas $\alpha_1, \ldots, \alpha_m$ y siendo $(b_k \pm ic_k$ $(1 \leq k \leq n)$ la *familia de las parejas de polos complejos conjugados* (no reales) de f, de órdenes respectivos $\beta_1, \ldots, \beta_n$. Al ser estos polos dos a dos distintos el conjunto de los polinomios de $\mathbf{R}[X]$ $(X - a_h)^{\alpha_h}$ $(1 \leq h \leq m)$, $[(X - b_k)^2 + (c_k)^2]^{\beta_k}$ $(1 \leq k \leq n)$ son *extraños dos a dos,* el teorema del § 205, *a*) nos permite escribir de una manera única

$$(6) \qquad f(X) = p(X) + \sum_{h=1}^{m} \sum_{j=1}^{\alpha_h} \frac{A_{hj}}{(X - a_h)^j} + \sum_{k=1}^{n} \sum_{l=1}^{\beta_k} \frac{B_{kl}X + C_{kl}}{[(X - b_k)^2 + (c_k)^2]^l}$$

siendo p un elemento de $\mathbf{R}[X]$ y reales todos los elementos

$$A_{hj}, \; B_{kl}, \; C_{kl} (1 \leq h \leq m, \qquad 1 \leq j \leq \alpha_h, \qquad 1 \leq k \leq n, \; 1 \leq l \leq \beta_k).$$

La parte relativa a un *polo real* a de orden α es siempre

$$P\left(\frac{1}{X - a} \right) = \frac{A_\alpha}{(X - a)^\alpha} + \ldots + \frac{A_1}{X - a}$$

polinomio real en $\frac{1}{X-a}$ de grado α y sin término constante; $P\left(\frac{1}{X-a}\right)$ es la suma e elementos simples de *primera especie.*

La parte relativa a una pareja de *polos conjugados* $b \pm ic$ (no reales) de orden β es la fracción real

$$R(X) = \frac{B_\beta^n X + C_\beta}{[(X-b)^2+c^2]^\beta} + \dots + \frac{B_1 X + C_1}{(X-b)^2+c^2}$$

donde $B_1, \dots, B_\beta, C_1, \dots, C_\beta$ son números reales, que pueden ser algunos nulos, aunque el polinomio $B_\beta X + C_\beta$ es no nulo, sino $b \pm ic$ serían polos de orden estrictamente inferior a β, lo que es imposible; $R(X)$ es la suma de elementos simples de *segunda especie* $\frac{B_l X + C_l}{[(X-b)^2+c^2]^l}$.

b) **Métodos prácticos de descomposición en R(X)**

El cálculo de p (parte entera) y de los coeficientes relativos a los polos reales se hace con los mismos métodos que en C(X) (ver § 206).

Para calcular las partes relativas a los polos complejos conjugados no reales la mayoría de las consideraciones desarrolladas en el § 206 son válidas. Por ejemplo, si

$$f(X) = \frac{u(X)}{[(X-b)^2+c^2]v_1(X)} = \sum_{l=1}^{\beta} \frac{B_l X + C_l}{[(X-b)^2+c^2]^l} + \frac{u_1(X)}{v_1(X)} \quad (v_1(b \pm ic) \neq 0)$$

se tendrá

$$\frac{u(X)}{v_1(X)} = B_\beta X + C_\beta + [(X-b)^2+c^2]k(X)$$

donde la fracción k está desprovista de los polos $b \pm ic$, de donde sustituyendo X por $b+ic$ la relación compleja

$$\frac{u(b+ic)}{v_1(b+ic)} = B_\beta(b+ic) + C_\beta$$

que determinan los dos números reales B_β y C_β.

Igualmente todas las observaciones relativas a la paridad o imparidad son válidas, así como el uso de valores particulares; respecto a esto hay que señalar que si $X^2 + pX + q$ es un polinomio real irreducible se tiene $q > 0$, entonces hay, a menudo, interés en utilizar

$$x = \pm i\sqrt{q}.$$

Se puede también observar que si se descompone f *real* en C(X) (en elementos simples de *primera especie*), las partes relativas a dos polos conjugados son conjugadas (§ 206, *c*); tendremos, pues, según la unicidad de la des-

composición (5) y de la descomposición (6), todos los demás términos de las dos descomposiciones estando desprovistos de polos $b \pm ic$

$$R(X) = \sum_{l=1}^{\beta} \frac{B_l X + C_l}{[(X-b)^2 + c^2]^l} = \sum_{l=1}^{\beta} \left[\frac{A_l}{(X-b-ic)^l} + \frac{\bar{A}_l}{(X-b+ic)^l} \right]$$

$$= \sum_{l=1}^{\beta} \frac{h_l(X)}{[(X-b)^2 + c^2]^l}$$

con $h_l(X) = A_l(X-b-ic)^l + \bar{A}_l(X-b-ic)^l \in \mathbf{R}[X]$. Puede resultar cómodo de descomponer en elementos simples de segunda especie, cada una de las fracciones reales $\dfrac{h_l(X)}{[(X-b)^2 + c^2]^l}$ utilizando, por ejemplo, el método de la división indicada en el ejercicio 2 del § 204.

OBSERVACION

De hecho, en la práctica, la descomposición en elementos simples de segunda especie es principalmente útil para calcular las primitivas de las funciones racionales reales. Ahora bien, las fórmulas de derivación en $\mathbf{C}(X)$ (§ 202) y la definición de la primitiva de una función compleja de variable real (ver curso de Análisis) demuestran que, si $l > 1$, la función real definida sobre $\mathbf{R}$

$$x \to \frac{A_l}{(x-a)^l} + \frac{\bar{A}_l}{(x-\bar{a})^l} \qquad (l > 1, \quad a = b + ic \notin \mathbf{R})$$

tiene como primitiva la función real definida sobre $\mathbf{R}$

$$x \to \frac{1}{1-l} \left(\frac{A_l}{(x-a)^{l-1}} + \frac{\bar{A}_l}{(x-\bar{a})^{l-1}} \right).$$

Solamente se necesitan los elementos simples de segunda especie en el caso de los *polos simples* y en aquel en el que los *residuos de los polos múltiples* $b \pm ic$ *son no nulos*; se escribirá entonces

$$\frac{A_1}{x-b-ic} + \frac{\bar{A}_1}{x-b+ic} = \frac{B_1 x + C_1}{(x-b)^2 + c^2} \qquad (B_1,\ C_1 \in \mathbf{R}).$$

Se observará que el cálculo indicado para una primitiva, en el caso $l > 1$, es más simple que el cálculo de la primitiva de

$$x \to \frac{B_l x + C_l}{[(x-b)^2 + c^2]^l}.$$

En consecuencia, resulta interesante, para calcular una primitiva de una función racional real, descomponerla en $\mathbf{C}(X)$ *en elementos simples de primera especie, y seguir el procedimiento indicado más arriba.*

Por otra parte, como se verá en Análisis, para los cálculos relativos a los operadores 1, 2 y 3 (§ 205, *b*, observación), los elementos simples de segunda especie, en general, no son de utilidad; de ahí el escaso interés práctico de la descomposición en elementos simples de primera y segunda especie en $\mathbf{R}(X)$.

EJEMPLOS

1. Los ejemplos 1 y 2 del § 206, dan inmediatamente la parte relativa a la pareja única de polos complejos conjugados bajo la forma de elemento simple de segunda especie.

2.
$$f(X) = \frac{1}{(X+1)^3 (X^2+X+1)^2}$$

poniendo $X + 1 = Y$ se halla

$$f(X) = \frac{1}{(X+1)^3} + \frac{2}{(X+1)^2} + \frac{1}{X+1} + f_1(X)$$
$$f_1(X) = -\frac{X^3 + 3X^2 + 3X + 3}{(X^2+X+1)^2}$$

dividiendo $-(X^3 + 3X^2 + 3X + 3)$ por $X^2 + X + 1$ después el cociente por $X^2 + X + 1$ (ver § 204, ej. 2) se obtiene

$$f_1(X) = \frac{-1}{(X^2+X+1)^2} - \frac{X+2}{X^2+X+1}.$$

Se hubiera podido poner *a priori* $(X^2 + X + 1 = (X - j)(X - j^2))$

$$f(X) = \frac{A_3}{(X+1)^3} + \frac{A_2}{(X+1)^2} + \frac{A_1}{X+1} + \frac{B_2X + C_2}{(X^2+X+1)^2} + \frac{B_1X + C_1}{X^2+X+1}$$

se tendrá (después de la multiplicación por $(X+1)^3$) sustituyendo X por -1 seguido (después de la multiplicación por $(X^2 + X + 1)^2$) de la sustitución de X por j

$$A_3 = 1, \qquad \frac{1}{(j+1)^3} = -1 = B_2 j + C_2 \Rightarrow (B_2 = 0, \quad C_2 = -1)$$

se podrá seguidamente dar los valores particulares 0, i e ∞ (después de la multiplicación por X), se obtendrá cuatro relaciones reales que permiten calcular A_2, A_1, B_1, C_1 (se constatará que este cálculo es más largo que el precedente).

3. $$f(X) = \frac{8}{(X^2-1)(X^2+1)^2} = \frac{A}{X-1} + \frac{A'}{X+1} + \frac{B_2X + C_2}{(X^2+1)^2} + \frac{B_1X + C_1}{X^2+1}.$$

La paridad nos permite escribir

$$f(X) = \frac{A}{X-1} - \frac{A}{X+1} + \frac{C_2}{(X^2+1)^2} + \frac{C_1}{X^2+1}$$

multiplicando por $X - 1$ y sustituyendo X por 1 se obtiene $A = 1$; multiplicando por $(X^2 + 1)^2$ y sustituyendo X por i se obtiene $C_2 = -4$; finalmente sustituyendo X por 0 se obtiene

$$-8 = -2A + C_2 + C_1 \Rightarrow C_1 = -2.$$

Ejercicios

369. Descomponer en elementos simples sobre **C** después sobre **R** ($n \in \mathbf{N}$)

$$\frac{1}{X^{2n}-1}, \quad \frac{1}{X^{2n+1}-1}, \quad \frac{1}{X^{2n}+1}, \quad \frac{1}{X^{2n+1}+1}.$$

370. Descomponer en elementos simples sobre **C** después sobre **R** ($p,\ q \in \mathbf{N},\ p \leq 2q$)

$$\frac{X^{p-1}}{1-X^{2q}}, \quad \frac{X^{p}}{1-X^{2q+1}}, \quad \frac{X^{p-1}}{1+X^{2q}}, \quad \frac{X^{p}}{1+X^{2q+1}}.$$

371. Descomponer en elementos simples sobre **C** ($n \in \mathbf{N};\ a,\ b \in \mathbf{C},\ a \neq b$)

$$\frac{1}{(X^2-1)^n}, \quad \frac{1}{(X-a)^n(X-b)^n}.$$

(Se podrá utilizar el método indicado al final del subpárrafo 206, **b**).

372. Descomponer en elementos simples sobre **C** después sobre **R**

$$\frac{1}{(X-1)^{10}(X^2+2X\times 4)}, \quad \frac{5X^6-X^3+1}{X^4+2X^3-6X^2+2X+1}, \quad \frac{3X^2+1}{(X-1)^4(X^2-X+1)^2},$$

$$\frac{X^4+2}{X(X^2-1)^2(X^2+1)}, \quad \frac{X^2-3}{X^2(X^2-1)(X^2+1)^3}, \quad \frac{X^5-2X^3+1}{X^4-2X^3+2X^2-2X+1}.$$

373. Descomponer en elementos simples sobre **C** después sobre **R** ($a,\ b \in \mathbf{R},\ a^2 \neq b^2$)

$$\frac{1}{X^4-2X^3(\cos a+\cos b)+2X^2(1+2\cos a\cos b)-2X(\cos a+\cos b)+1}.$$

374. Descomponer en elementos simples sobre **C** seguidamente sobre **R** (eventualmente)

$$\frac{1}{4X^3-3X-a}$$

($a \in \mathbf{R}$; se pondrá $a = \cos 3\alpha$ o $a = ch\ 3\alpha$ o $a = -ch\ 3\alpha$ según las posiciones de a con relación a -1 y 1).

375. Tenemos la fracción racional f definida por

$$x = \operatorname{tg} a, \qquad f(x) = \operatorname{tg} na, \qquad \left(a \in \mathbf{R}, -\frac{\pi}{2} < a < \frac{\pi}{2} \right)$$

Descomponer f en elementos simples.

376. Sea $u = ax^2 + bx + c$, $v = a'X^2 + b'X + c'$ dos trinomios con coeficientes reales; se supone $a'(b'^2 - 4a'c') \neq 0$. Descomponer $f = u/v$ en elementos simples de primera especie. Calcular f''. Deducir que f'' tiene 1 o 3 raíces según que v tenga 2 o 0 raíces reales.

377. **a)** Dados dos trinomios de segundo grado T_1, T_2, primos entre sí, con coeficientes complejos, demostrar que las raíces de los trinomios $T = \alpha_1 T_1 + \alpha_2 T_2$, en los que α_1 y α_2 son dos números complejos cualesquiera, verifican una relación involutiva (V. § 124, f) que no depende más que de T_1 y T_2.

b) ¿En qué condición los residuos de la fracción

$$\frac{(X-\alpha)(X-\beta)}{(X-a)^2(X-b)^2}$$

son nulos? (α, β, a, $b \in \mathbf{C}$; $a \neq b$).

378. Siendo (x_h) $(1 \leq h \leq n)$ una familia de n números complejos todos distintos, se pone

$$u(X) = (X-x_1)(X-x_2)\ \ldots\ (X-x_n).$$

Demostrar que

$$\frac{1}{[u(X)]^2} = \sum_{h=1}^{n}\left[\frac{A_h}{(X-x_h)^2} + \frac{B_h}{X-x_h}\right]$$

Calcular A_h y B_h con ayuda de $u'(x_h)$ y $u''(x_h)$

379. Siendo (x_h) $(1 \leq h \leq n)$ una familia de n números complejos todos distintos y (a_h) $(1 \leq h \leq n)$ una familia de números complejos cualesquiera, demostrar que existe un polinomio único f de $\mathbf{C}[X]$, de grado $\leq n-1$ tal que

$$(h = 1, 2, \ldots, n) \qquad f(x_h) = a_h$$

considerando la fracción

$$\frac{f(X)}{(X-x_1)\ \ldots\ (X-x_n)}.$$

Hallar de nuevo la fórmula de interpolación de Lagrange (§ 192, ej. 5).

380. Siendo a y b dos números complejos distintos y α, α', β, β' cuatro números complejos cualesquiera, demostrar que hay un polinomio único f de $\mathbf{C}[X]$ de grado ≤ 3 tal que

$$f(a) = \alpha, \qquad f'(a) = \alpha', \qquad f(b) = \beta, \qquad f'(b) = \beta'.$$

(Considerar la fracción $f(X)/(X-a)^2(X-b)^2$).

381*. Si (x_h) $(1 \leq h \leq m)$ es una familia de m números complejos todos distintos, (p_h) $(1 \leq h \leq m)$ una familia de enteros naturales no nulos y $(a_{h,k})$ $(0 \leq k \leq p_h - 1)$ una familia de números complejos cualesquiera, estudiar la existencia y unicidad de un polinomio f de $\mathbf{C}[X]$ de grado $\leq p_1 + \ldots + p_m - 1$ tal que

$$(h = 1, \ldots, m) \qquad f(x_h) = a_{h,0}\ \ldots\ f^k(x_h) = a_{h,k}\ \ldots\ f^{(p_h-1)}(x_h) = a_{h,p_h-1}.$$

(Estudiar primero los ejercicios 379 y 350.)

382. Demostrar que en $\mathbf{C}(X)$

$$\frac{1}{1-X} = 1 + X + \ldots + X^n + \frac{X^{n+1}}{1-X}.$$

Deducir por derivación los cocientes del orden n en la división según las potencias crecientes de 1 por $f(X) = (1-X)^k$.
Aplicar el resultado a $f(X) = (1-aX)^k$ o $f(X) = (a-X)^k$ $(a \in \mathbf{C}^*)$.

383. Descomponer sobre $\mathbf{C}$, $f(X) = 1/(1+X^2)$. Deducir que

$$f^n(X) = \frac{P_n(X)}{(1+X^2)^n}.$$

donde P_n es un polinomio de $\mathbf{R}[X]$ de grado n que se calculará.

384. Los polinomios considerados se supone que tienen sus coeficientes en un cuerpo conmutativo **K**.
Demostrar que si $\nu(x)$ es un polinomio no nulo, todo polinomio f se escribe de una manera única

$$f(X) = \Sigma a_h(X)\,[\nu(X)]^h,$$

los polinomios a_h son de grado $<$ grd ν cuando no son nulos.
(Efectuar en $\mathbf{K}[Y][X]$ la división euclídea de $f(X)$ por $\nu(X) - Y$).
Deducir del resultado precedente una nueva demostración del lema 3 (§ 204).

385. Descomponer en elementos simples sobre **R**

$$\frac{3X^7 - 5X^4 + 4X^2 - 11X + 1}{(X^2 + X + 1)^{1000}},$$

(utilizar el ejercicio 384).

386. 1.° Se pone $f(X) = \dfrac{1}{X^3 - 1}$. Calcular la parte principal

$$\frac{A}{(X-1)^3} + \frac{B}{(X-1)^2} + \frac{C}{X-1} \qquad \text{de} \qquad [f(X)]^3 = g(X)$$

relativa al polo 1.
2.° Observando que $g(X) = g(jX) = g(j^2X)$ $(j = e^{2i\pi/3})$ hallar la descomposición en elementos simples de g en $\mathbf{C}(X)$.
3.° Demostrar que existe un valor de λ tal que $g(X) - \lambda f(X)$ sea la derivada de una fracción racional $G(X)$. Suponiendo $G(0) = 0$ escribir la descomposición de $G(X)$ en $\mathbf{C}(X)$.
4.° Se pone $\varphi(X) = 1/(X^n - 1)$, n entero ≥ 2. Demostrar que existe un valor λ para el cual la fracción

$$[\varphi(X)]^3 - \lambda\varphi(X)$$

es la derivada de una fracción racional. Calcular λ. (M.P.C.)

387. Se tiene la fracción racional $(a,\ b \in \mathbf{C})$

$$f(X) = \frac{aX^2 + bX + c}{(X-1)^2\,(X+1)^2}.$$

a) Descomponer f en elementos simples.
b) ¿En qué condición las primitivas de f son fracciones racionales en X? Dar su expresión en este caso.
(V. ej. 377). (M.G.P.)

388. Hallar los polinomios $u(X)$ de grado mínimo tal que las primitivas de la fracción racional

$$f(X) = \frac{u(X)}{X^3(X^4 + 1)^2}$$

sean fracciones racionales en X.

389. Sea E el espacio vectorial de los polinomios con coeficientes complejos E_{n+1} el subespacio de los polinomios de grado $\leq n$. En todo lo que sigue Q es un polinomio dado de grado q en el que todas las raíces son simples.
1.° Si P es un polinomio, ¿cuál es la forma de la descomposición de $z = P/Q^2$ en elementos simples? Demostrar que si las primitivas de z son racionales, son de la forma S/Q donde S es un polinomio.

2.º Se hace corresponder a $U \in E_{n+1}$ el polinomio

$$T(U) = U'Q - Q'U$$

demostrar que T es una aplicación lineal de E_{n+1} en E_{n+q}. ¿Cuál es el núcleo de T? (Se distinguirá el caso $n \geq q$ y $n < q$.) ¿Cuál es el rango de T, la dimensión de la imagen $T(E_{n+1})$ de T?
3.º Siendo U exactamente de grado n, demostrar que si T(U) es de grado $< n + q - 1$ entonces $n = q$ y que existe un polinomio V de grado $< n$ tal que $T(V) = T(U)$.
Dado el entero natural p, deducir que los polinomios $P \in E_{p+1}$, tales que la fracción P/Q^2 tenga sus primitivas racionales formen un espacio vectorial de dimensiones $p - q + 1$, si $p \geq 2q - 1$, $p - q + 2$ si $q - 2 \leq p < 2q - 1$, y cero si $p < q - 2$.
Verificar estos resultados sobre el ejemplo estudiado en el ejercicio 387.

(M.G.P.)

390. Sea (x_h) $(1 \leq h \leq m)$ una familia de m números complejos todos distintos.
a) Si a_0 es un número complejo no nulo, calcular la descomposición sobre **C** de la $f'(X)/f(X)$ con $f(X) = a_0(X - x_1)^{k1} \ldots (X - \alpha_m)^{km} \in \mathbf{C}[X]$.
b) Demostrar que para que exista un polinomio $g \neq 0$ de $\mathbf{C}[X]$, tal que

$$\frac{g'(X)}{g(X)} = \sum_{h=1}^{m} \frac{\alpha_h}{X - x_h}$$

es necesario y suficiente que $\alpha_1, \ldots, \alpha_m$ sean enteros naturales. Hallar entonces todos los polinomios g.

391. Simplificar, para $n = 2, 3, 4$ las fracciones racionales (V. ej. 366, cap. 11)

$$\frac{X^n}{(X-Y)(X-Z)} + \frac{Y^n}{(Y-Z)(Y-X)} + \frac{Z^n}{(Z-X)(Z-Y)}.$$

392 Simplificar, para $n = 2, 3, 4$, las fracciones

$$X^n \frac{(X+Y)(X+Z)}{(X-Y)(X-Z)} + Y^n \frac{(Y+Z)(Y+X)}{(Y-Z)(Y-X)} + Z^n \frac{(Z+X)(Z+Y)}{(Z-X)(Z-Y)}.$$

393. Hallar los ceros, los polos y los puntos de indeterminación de las fracciones siguientes de $\mathbf{C}(X, Y)$ o $\mathbf{C}(X, Y, Z)$

$$\frac{X+Y}{X-Y} \qquad \frac{X^3+Y^3}{X^2-Y^2}, \qquad \frac{X+Y+Z}{X-Y}.$$

394*. **a)** Hallar los ceros comunes de cada uno de los pares de polinomios formados con los tres polinomios de $\mathbf{C}[X, Y, Z, T]$

$$Y^2 - XZ, \qquad YT - X^2, \qquad XY - ZT.$$

b) Hallar los ceros, los polos, los puntos de indeterminación de las tres fracciones de $\mathbf{C}(X, Y, Z, T)$

$$\frac{Y^2 - XZ}{YT - X^2} \qquad \frac{YT - X^2}{XY - ZT} \qquad \frac{XY - ZT}{Y^2 - XZ}.$$

Se demostrará, en particular, que estas tres fracciones tienen como puntos de indeterminación todos los puntos de $\mathbf{C}^4$ de la forma $(\lambda t, \lambda t^2, \lambda t^3, \lambda)$ en que en que λ y t son dos números complejos cualesquiera.

13

ECUACIONES ALGEBRAICAS

208. Introducción

En todo este capítulo sólo consideraremos polinomios de C[X]. Siendo C infinito, a todo polinomio f de C[X] podemos hacerle corresponder, de manera biyectiva, la aplicación polinómica asociada (ver § 185), representaremos igualmente esta aplicación por f. Por otra parte, el polinomio $f \neq 0$ de grado n

(1) $$f(X) = a_0X^n + a_1X^{n-1} + \ldots + a_{n-1}X + a_n \qquad (a_0 \neq 0)$$

tiene n raíces distintas o no en C (ver § 193, b); según que expresemos o no de modo explícito el orden de multiplicidad de cada raíz escribiremos (2) o (2′)

(2) $$f(X) = a_0(X - \alpha_1) \ldots (X - \alpha_n)$$

(2′) $$f(X) = a_0(X - \alpha_1)^{h_1} \ldots (X - \alpha_m)^{h_m}.$$

Donde las m raíces $\alpha_1, \ldots, \alpha_m$ de órdenes respectivos $h_1, \ldots, h_m$ son distintos dos a dos; además, $h_1 + \ldots + h_m = n$.

Hallar las raíces del polinomio f, es *resolver la ecuación* $f(x) = y$, asociada a f (ver § 14, b) para $y = 0$. Diremos que la ecuación

(3) $$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_{n-1}x + a_n = 0 \qquad (a_0 \neq 0)$$

es una *ecuación algebraica de grado n* (se sobrentiende: con coeficientes $a_0, \ldots, a_n$ en C).

Las raíces del polinomio f se llaman *raíces* o soluciones de la ecuación $f(x) = 0$. Dos ecuaciones algebraicas $f(x) = 0$, $g(x) = 0$ son *equivalentes* si tienen las mismas raíces con el mismo orden de multiplicidad: esto ocurre si y sólo si $g = \lambda f$ ($\lambda \in C^*$).

El lector se dará cuenta fácilmente que los resultados que vamos a dar serán también válidos si se reemplaza C por un cuerpo K, conmutativo, infinito, algebraicamente cerrado, de característica nula. Algunos de ellos son también válidos en casos más generales.

I. Funciones racionales de las raíces

209. Relación entre los coeficientes y las raíces

Sea $\alpha_1, \ldots, \alpha_n$ las n raíces de la ecuación algebraica de grado n, $f(x) = 0$, si $\Sigma_1, \ldots, \Sigma_n$ son los polinomios simétricos elementales de $\mathbf{C}[X_1, \ldots, X_n]$ (ver § 199, *b*) escribiremos

$$(h = 1, \ldots, n) \qquad \Sigma_h(\alpha_1, \ldots, \alpha_n) = \Sigma\alpha_1\alpha_2 \ldots \alpha_h = \sigma_h$$

y diremos que $\sigma_1, \ldots, \sigma_n$ son las *n funciones simétricas elementales de las raíces de* $f(x) = 0$.

Tenemos

$$(4) \qquad \begin{cases} \sigma_1 = \Sigma\alpha_1 = \alpha_1 + \alpha_2 + \ldots + \alpha_n \\ \sigma_2 = \Sigma\alpha_1\alpha_2 = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \ldots + \alpha_{n-1}\alpha_n \\ \vdots \\ \sigma_h = \Sigma\alpha_1\alpha_2 \ldots \alpha_h = \alpha_1\alpha_2 \ldots \alpha_h + \ldots + \alpha_{n-h+1} \ldots \alpha_n \\ \vdots \\ \sigma_n = \Sigma\alpha_1\alpha_2 \ldots \alpha_n = \alpha_1\alpha_2 \ldots \alpha_n. \end{cases}$$

Las fórmulas (1) y (2) (§ 208) y los cálculos del § 199, *b*) demuestran que (si f es de grado n y $a_0 \neq 0$)

$$(5) \qquad \boxed{\sigma_1 = -\frac{a_1}{a_0}, \ \ldots, \ \sigma_h = (-1)^h \frac{a_h}{a_0}, \ \ldots, \ \sigma_n = (-1)^n \frac{a_n}{a_0}}$$

Recíprocamente supongamos que, siendo $\sigma_1, \ldots, \sigma_n$ n números complejos dados, tratásemos de resolver el sistema (4), $\alpha_1, \ldots, \alpha_n$ como incógnitas, está claro que $\alpha_1, \ldots, \alpha_n$ son las n raíces de la ecuación

$$(6) \qquad x^n - \sigma_1 x^{n-1} + \ldots + (-1)^h \sigma_h x^{n-h} + \ldots + (-1)^n \sigma_n = 0.$$

EJEMPLOS Y EJERCICIOS

1. Escribir las relaciones (5) separando una raíz α_1 e introduciendo las $n-1$ funciones simétricas elementales $\sigma'_1, \ldots, \sigma'_{n-1}$ de $\alpha_2, \alpha_3, \ldots, \alpha_n$; se obtiene

$$(5') \qquad \begin{cases} \alpha_1 + \sigma'_1 = -a_1/a_0 \\ \vdots \\ \alpha_1\sigma'_{h-1} + \sigma'_h = (-1)^h a_h/a_0 \qquad (h = 2, \ldots, n-1) \\ \vdots \\ \alpha_1\sigma'_{n-1} = (-1)^n a_n/a_0 \end{cases}$$

deducir que el sistema (5) en el que $\alpha_1, \ldots, \alpha_n$ son las incógnitas es equivalente al sistema (5′)

$$f(\alpha_1) = f(\alpha_2) = \ldots = f(\alpha_n) = 0.$$

Demostrar que si en el sistema (5) se reemplazan $m < n$ ecuaciones por m ecuaciones del sistema (5′) se obtiene una consecuencia del sistema (5). Así (dados a, b, c) se tiene

$$\begin{cases} x + y + z = -a \\ yz + zx + xy = b \\ xyz = -c \end{cases} \Rightarrow \begin{cases} x + y + z = -a \\ xyz = -c \\ x^3 + ax^2 + bx + c = 0. \end{cases}$$

¿En qué caso estos dos últimos sistemas son equivalentes?

2. ¿En qué condición las raíces de $x^3 + ax^2 + bx + c = 0$ están en progresión aritmética? Resolver entonces la ecuación.

3. Para $n = 4$ escribir las fórmulas (5) poniendo

$$s = \alpha_1 + \alpha_2, \qquad p = \alpha_1\alpha_2, \qquad s' = \alpha_3 + \alpha_4, \qquad p' = \alpha_3\alpha_4.$$

4. Hallar a tal que las raíces de $x^4 + x^3 + ax^2 + 3x + 2 = 0$ verifiquen $\alpha_1 + \alpha_2 = \alpha_3\alpha_4$. Resolver entonces la o las ecuaciones obtenidas.

5. Utilizando σ_1 hallar h de modo que haciendo el cambio de variable $x = y + h$ la ecuación (3) tome la forma

$$(3') \qquad y^n + b_2y^{n-2} + b_3y^{n-3} + \dots + b_{n-1}y + b_n = 0.$$

Así para $n = 3$, se tiene $y^3 + py + q = 0$, algunas veces se dice que se tiene la «forma canónica» de la ecuación de tercer grado.

210. Funciones racionales de las raíces de una ecuación algebraica

Sea r una fracción racional simétrica, elemento de $\mathbf{C}(X_1, \dots, X_n)$. Para toda permutación π de $[1, n]$ se tiene

$$(1) \qquad r(X_{\pi(1)}, \dots, X_{\pi(n)}) = r(X_1, \dots, X_n)$$

resultará, supuestas *sustituibles* en r las raíces $\alpha_1, \dots, \alpha_n$ de $f(x)$, que para toda permutación π

$$(2) \qquad r(\alpha_{\pi(1)}, \dots, \alpha_{\pi(n)}) = r(\alpha_1, \dots, \alpha_n).$$

Pero puede suceder que se tenga (2) sin que se tenga (1). Consideremos, por ejemplo,

$$f(x) = x^3 + px + q = (x - a)(x - b)(x - c)$$

tenemos $(a + b + c = 0)$

$$a^2 - bc = a(a + b + c) - (bc + ca + ab) = -(bc + ca + ab)$$

luego

$$a^2 - bc = a^2 - cb = b^2 - ca = b^2 - ac = c^2 - ab = c^2 - ba = -\sigma_2 = -p$$

aunque el polinomio $X^2 - YZ$ no sea simétrico.

Si una función racional r que aplica $\mathbf{C}^n$ en $\mathbf{C}$ verifica (2) para las n raíces de una ecuación $f(x) = 0$, para toda permutación π de $[1, n]$, diremos que r es *numéricamente simétrica* para las n raíces $\alpha_1, \dots, \alpha_n$ de $f(x) = 0$; por abuso de lenguaje, diremos que r es una *función racional simétrica de las n raíces de $f(x) = 0$*. Por oposición, se dice algunas veces que s elemento simétrico de $\mathbf{C}(X_1, \dots, X_n)$ es *formalmente simétrico.*

Supongamos **r numéricamente simétrico** para las raíces de $f(x) = 0$, si efectuando sobre r *todas* las permutaciones de $\mathcal{S}_n$, se obtiene k y solamente k fracciones racionales $r_1, \ldots, r_k$ distintas dos a dos es evidente que la fracción racional

$$s(X_1, \ldots, X_n) = \frac{1}{k}[r_1(X_1, \ldots, X_n) + \ldots + r_k(X_1, \ldots, X_n)]$$

es formalmente simétrica y que

$$s(\alpha_1, \ldots, \alpha_n) = r(\alpha_1, \ldots, \alpha_n)$$

Por ejemplo, en el caso estudiado anteriormente

$$s(X, Y, Z) = \frac{1}{3}(X^2 - YZ + Y^2 - ZX + Z^2 - XY)$$

y

$$s(a, b, c) = r(a, b, c).$$

En consecuencia: el cálculo de una *función racional numéricamente simétrica* de las raíces de una ecuación algebraica, se reduce al cálculo del valor de una *fracción racional formalmente simétrica* para las raíces de la ecuación estudiada.

El teorema 21 del § 199 y las fórmulas (5) del § 209 nos permiten enunciar:

TEOREMA. — *Toda fracción racional numéricamente simétrica de las raíces de una ecuación algebraica, se expresa racionalmente en función de los coeficientes de esta ecuación.*

EJERCICIOS

1. Sea r una fracción racional simétrica, de representante irreducible P/Q, P y Q siendo dos polinomios simétricos de grados parciales respectivos p' y q'. Dar la expresión de $r(\alpha_1, \ldots, \alpha_n)$ en función de los coeficientes $a_0, \ldots, a_n$ de la ecuación que tiene por raíces $\alpha_1, \ldots, \alpha_n$. ¿Qué se hallará si $r = P$, P polinomio simétrico de grado total p y de grado parcial p'?
2. Sea $f(x) = x^3 + px + q = (x - a)(x - b)(x - c)$, calcular $(b - c)^2(c - a)^2(a - b)^2$ en función de p y q, α) directamente, β) utilizando el teorema del grado y del peso (§ 199) (ver igualmente un tercer método, ej. 424, b).
3. Sea $f(x) = x^3 + px + q = (x - a)(x - b)(x - c)$, con $p + q + 1 \neq 0$, calcular

$$\sum \frac{a - 1}{b - 1}.$$

II. Eliminación y aplicaciones

211. Definición. Métodos teóricos de eliminación

DEFINICIÓN. — *Sean dos ecuaciones algebraicas*

$$f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n = 0 \qquad (a_0 \neq 0)$$
$$g(x) = b_0x^p + b_1x^{p-1} + \ldots + b_p = 0 \qquad (b_0 \neq 0)$$

eliminar x *entre estas dos ecuaciones, es hallar una condición que la verifiquen los coeficientes de las dos ecuaciones y sea, además, necesaria y suficiente para que estas dos ecuaciones tengan al menos una raíz común.*

Designaremos las raíces de $f(x)=0$ por $\alpha_1, \ldots, \alpha_n$ y las de $g(x)$ por $\beta_1, \ldots, \beta_p$. Escribiremos $n \leq p$.

a) Método del máximo común divisor

Si los polinomios f y g tienen una raíz común γ, son divisibles por $X-\gamma$ y al ser γ una raíz de su máximo común divisor d, este último es de grado superior o igual a 1. Recíprocamente si el m. c. d. d de los polinomios f y g es de grado superior o igual a 1, los polinomios f y g tienen en común las raíces de d.

Buscando el m. c. d. de f y g por el algoritmo de EUCLIDES (§ 190, *c*) se obtendrá r_k tal que

$$\text{grd } r_{k-1} > 0 \qquad \text{grd } r_k = 0$$

para que f y g tengan una raíz común es necesario y suficiente que $r_k = 0$; en general r_{k-1} será de primer grado, habrá una raíz común única, si r_{k-1} es de grado l, f y g y, por tanto, las ecuaciones $f(x)=0$ y $g(x)=0$ tendrán l raíces comunes.

EJEMPLOS Y EJERCICIOS

1. Si grd $f=n$ y grd $g=1$, la raíz común no puede ser otra que la raíz $-b_1/b_0$ de g, la condición es

$$(b_0)^n f\left(-\frac{b_1}{b_0}\right) = a_0(-b_1)^n + a_0 b_0(-b_1)^{n-1} + \ldots + a_n(b_0)^n = 0.$$

2. Sea $f(x) = ax^2+bx+c$, $g(x) = a'x^2+b'x+c'$ $(aa' \neq 0)$; demostrar que f y g tienen al menos una raíz común si y sólo si

$$\mathrm{R} = (ca'-ac')^2 - (ab'-ba')(bc'-cb') = 0.$$

Demostrar que si $\mathrm{R}=0$ y $ab'-ba'=0$, f y g tienen dos raíces comunes.

¿En qué caso f y g tienen una raíz única común? Hallar en este caso esta raíz.

3. Si a y c son no nulos $a+bx+cx^2=0$ tiene por raíces las inversas de las raíces de $ax^2+bx+c=0$. Situándonos en $\widetilde{\mathbf{C}} = \mathbf{C} \cup \{\infty\}$ (ver § 124) diremos que la ecuación $ax^2+bx+c=0$ admite ∞ por raíz simple si $a=0$, $b\neq 0$ y admite ∞ por raíz doble si $a=b=0$, $c\neq 0$.

Si a y a' pueden ser nulos, demostrar que la condición $\mathrm{R}=0$ (ver ejercicio 2 más arriba) es necesaria y suficiente para que $f(x)=0$ y $g(x)=0$ tengan al menos una raíz común finita o infinita.

b) Determinante de Sylvester

f y g tienen una raíz común si y sólo si su m. c. d. d es de grado $l>0$, es decir, si se tiene

(1) $\qquad f=f_1 d, \qquad g=g_1 d \qquad (\text{grd } f_1 < n, \qquad \text{grd } g_1 < p)$

existe, pues, f_1 y g_1 verificando

$$(2) \qquad g_1f - f_1g = 0 \qquad (\text{grd } f_1 < n, \quad \text{grd } g_1 < p).$$

Recíprocamente si existe f_1 y g_1 verificando (2), f y g no son extraños, pues si lo fueran f al dividir $fg_1 = f_1g$ dividiría f_1, lo que es imposible, de donde:

TEOREMA. — *Las ecuaciones algebraicas $f(x) = 0$, $g(x) = 0$ tienen una raíz común si y sólo si existen polinomios u y v no nulos que verifican*

$$(3) \qquad uf + vg = 0 \qquad (\text{grd } u < \text{grd } g, \quad \text{grd } v < \text{grd } f).$$

Si f y g tienen una raíz común existe, en consecuencia, números complejos $c_0, \dots, c_{p-1}$ no todos nulos y números complejos $d_0, \dots, d_{n-1}$ no todos nulos tales que

$$(4) \qquad (c_0X^{p-1} + \dots + c_{p-1})f + (d_0X^{n-1} + \dots + d_{n-1})g = 0.$$

Luego los $n + p$ polinomios

$$X^hf \qquad (0 \leq h \leq p-1), \qquad X^kg \qquad (0 \leq k \leq n-1)$$

describen una parte ligada del espacio vectorial E de los polinomios de grado a lo sumo igual a $n + p - 1$. Recíprocamente si estos polinomios no son independientes, existe una relación de forma (4) con coeficientes no todos nulos: está claro que todos los c_h (o todos los d_k) no pueden ser nulos, se tendría entonces $vg = 0$ con $v \neq 0$, $g \neq 0$, lo que es imposible; luego si estos $n + p$ polinomios no son independientes existe u y v no nulos, verificando (3).

La condición f y g tienen una raíz común, es equivalente a la condición los $n + p$ polinomios X^hf, X^kg $(0 \leq h \leq p-1,\ 0 \leq k \leq n-1)$ no son independientes: es decir, la matriz de sus coeficientes sobre la base canónica de E: $\{X^{n+p-1}, \dots, X, 1\}$ no es inversible, o sea, su determinante S es nulo

$$S = \overbrace{\qquad\qquad\qquad}^{p}\ \overbrace{\qquad\qquad}^{n}$$

$$S = \begin{vmatrix}
a_0 a_1 \dots a_{n-1} a_n \ 0 \dots 0 & 0 \dots\dots\dots 0 \\
0\ a_0 a_1 \dots\dots a_{n-1} a_n\ 0\ 0 & 0 \dots\dots\dots 0 \\
\vdots \qquad \vdots \qquad\qquad \vdots & \vdots \qquad\qquad \vdots \\
0\ 0 \ \dots\ 0 \ \dots\dots\dots a_0 & a_1 \dots\dots\dots a_n \\
b_0 b_1 \dots b_{n-1} \dots\dots\dots b_{p-1} b_p & 0 \dots\dots 0 \\
b_0 b_1 \dots\dots\dots\dots\dots\dots & b_{p-1} b_p\ 0 \dots 0 \\
\vdots \qquad \vdots \qquad\qquad \vdots & \vdots \qquad\qquad \vdots \\
0\ 0 \ \dots\ b_0 b_1 \dots\dots\dots b_{p-n} & b_{p-n+1} \dots\dots b_p
\end{vmatrix} = 0 \quad \begin{matrix} \uparrow \\ p \\ \downarrow \\ \uparrow \\ n \\ \downarrow \end{matrix}$$

El determinante S se puede considerar como un polinomio de $n + p + 2$ indeterminadas $a_0, \dots, a_n, b_0, \dots, b_p$ se le llama determinante de SYLVESTER de los dos polinomios f y g; tenemos, pues:

TEOREMA. — *Dos polinomios f y g tienen una raíz común si y sólo si su determinante de Sylvester es nulo.*

EJEMPLOS Y EJERCICIOS

4. Calcular el determinante de SYLVESTER para $f = ax^2 + bx + c$, $g = a'x^2 + b'x + c'$. Verificar que $R = S$ (ver ej. 2 más arriba).

5. Demostrar que S es un polinomio homogéneo de grado p relativamente a $a_0, \dots, a_n$ y un polinomio homogéneo de grado n relativamente a $b_0, \dots, b_p$.

6. Demostrar que $a_0 = b_0 = 0$ implica $S = 0$. Situándonos en $\tilde{C}$ y a_0 o b_0 pudiendo ser nulos, demostrar que la condición $S = 0$ es necesaria y suficiente para que las dos ecuaciones $f(x) = 0$ y $g(x) = 0$ tengan al menos una raíz común finita o infinita.

c) Métodos de funciones simétricas

Tenemos $(a_0 b_0 \neq 0)$

$$f(x) = a_0 x^n + \dots + a_n = a_0(x - \alpha_1) \dots (x - \alpha_n) = a_0 \prod_{i=1} (x - \alpha_i)$$

$$g(x) = b_0 x^p + \dots + b_p = b_0(x - \beta_1) \dots (x - \beta_p) = b_0 \prod_{j=1}^{p} (x - \beta_j).$$

Pongamos

$$A = \prod_{j=1}^{p} f(\beta_j) = (a_0)^p \prod_{i=1}^{n} \prod_{j=1}^{p} (\beta_j - \alpha_i)$$

$$B = \prod_{i=1}^{n} g(\alpha_i) = (b_0)^n \prod_{i=1}^{n} \prod_{j=1}^{p} (\alpha_i - \beta_j).$$

Se ve inmediatamente que

$$(1) \qquad R = (a_0)^p B = (-1)^{np}(b_0)^n A = (a_0)^p (b_0)^n \prod_{i=1}^{n} \prod_{j=1}^{p} (\alpha_i - \beta_j).$$

A es una función polinomio simétrica de las raíces de $g(x) = 0$; designemos por $\tau_1, \dots, \tau_p$ las funciones simétricas elementales de estas raíces; considerando $\beta_1, \dots, \beta_p$ como indeterminadas, siendo A de grado parcial n respecto de cada indeterminada β_j, tendremos (teorema del grado, § 199, *c*)

$$A = h_n(\tau_1, \dots, \tau_p)$$

donde h_n es un polinomio de grado total n cuyos coeficientes son polinomios en $a_0, \dots, a_n$, según la definición de A. Por otro lado, las relaciones entre los coeficientes y las raíces de $g(x) = 0$ (ver § 209) nos permiten escribir

$$A = h_n\left(-\frac{b_1}{b_0}, \dots, (-1)^p \frac{b_p}{b_0} \right) = \frac{P_n(b_0, \dots, b_p)}{(b_0)^n}$$

siendo P_n un polinomio homogéneo de grado total n relativamente a $b_0, \ldots, b_p$. Se demostraría igualmente que

$$B = \frac{Q_p(a_0, \ldots, a_n)}{(a_0)^p}$$

donde Q_p es un polinomio homogéneo de grado total p relativamente a $a_0, \ldots, a_n$.
Finalmente la fórmula (1) da

$$(2) \qquad R(a_0, \ldots, b_p) = Q_p(a_0, \ldots, a_n) = (-1)^{np} P_n(b_0, \ldots, b_p)$$

R *es, pues, un polinomio en* $a_0, \ldots, b_p$, *homogéneo de grado total* p *en* $a_0, \ldots, a_n$ *y homogéneo de grado total* n *en* $b_0, \ldots, b_p$, *se le llama el resultante de los dos polinomios* f *y* g.

Recordemos que los cálculos precedentes se han efectuado suponiendo $a_0 b_0 \neq 0$. En este caso $R = 0$ es equivalente a: existe (i, j) tal que $\alpha_i - \beta_j = 0$, de donde:

TEOREMA. — *Dos polinomios* f *y* g *tienen una raíz común si y sólo si su resultante es nulo.*

EJEMPLOS Y EJERCICIOS

7. Calcular R para $f(x) = ax^2 + bx + c$, $g(x) = a'x^2 + b'x + c'$. Comparar con los resultados los ejercicios 2 y 5 anteriores.

8. Demostrar que $R = S$ (considerar S como un polinomio en $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_p$ merced a las fórmulas $a_i = (-1)^i a_0 \sigma_i$, $b_j = (-1)^j b_0 \tau_j$ y observar que $S = 0$ si $\alpha_i = \beta_j$, utilizar seguidamente el corolario del teorema 18, § 195; comparar finalmente los grados de S y R en $a_0, \ldots, b_p$).

9. Demostrar que $R(a_0, \ldots, b_p)$ es isobaro y de peso np (ver § 199, c).

212. Métodos prácticos de eliminación

Los tres métodos de eliminación que acabamos de indicar en el párrafo precedente son a menudo aplicables en la práctica (ver ej. 2, 5 y 7, § 211). Se puede también observar que si existe dos polinomios f_1 y g_1 tales que, para x perteneciente a $\mathbf{C}$,

$$\begin{cases} f(x) = 0 \\ g(x) = 0 \end{cases} \Leftrightarrow \begin{cases} f_1(x) = 0 \\ g_1(x) = 0 \end{cases}$$

es equivalente eliminar x entre $f(x) = 0$ y $g(x) = 0$ o entre $f_1(x) = 0$ y $g_1(x) = 0$.

Pongamos $(n \leq p)$

$$(1) \qquad \begin{cases} f(x) = a_0 x^n + \ldots + a_n = 0 & (a_0 \neq 0) \\ g(x) = b_0 x^p + \ldots + b_p = 0 & (b_0 \neq 0) \end{cases}$$

el sistema (1) es equivalente al (2)

$$(2) \qquad \begin{cases} f(x) = 0 \\ g_1(x) = a_0 g(x) - b_0 x^{p-n} f(x) = 0. \end{cases}$$

El sistema (2) es más simple, pues grd $g_1 <$ grd g.

Igualmente si $a_n \neq 0$, (1) es equivalente a (3)

$$(3) \qquad \begin{cases} f(x) = 0 \\ g_2(x) = b_p f(x) - a_n g(x) = x g_3(x) = 0 \end{cases} \Leftrightarrow \begin{cases} f(x) = 0 \\ g_3(x) = 0 \end{cases}$$

pues $a_n \neq 0$ implica $f(0) \neq 0$. Ahora bien, grd $g_3 <$ grd g_2.

En cada uno de estos procedimientos se ha bajado al menos en una unidad el grado de una de las ecuaciones: se puede, pues, llegar poco a poco al caso en que una de las ecuaciones sea de primer grado.

Pero estos cálculos en la práctica requieren un gran cuidado: hay que pasar cada vez a un sistema equivalente. Ahora bien, en la práctica los coeficientes $a_0, \ldots, b_p$ dependen de parámetros y lo que se busca son valores de estos parámetros o las relaciones entre estos parámetros para que $f(x) = 0$ y $g(x) = 0$ tengan una raíz común. Supongamos para simplificar que $a_0, \ldots, b_p$ sean polinomios en λ, escribiendo que $f(x) = 0$ y $g_1(x) = 0$ (sistema (2) anterior) tienen una raíz común, se encontrará una condición $\varphi(\lambda) = 0$. Si una solución λ_0 de $\varphi(\lambda) = 0$ es tal que $a_0(\lambda) = 0$ los sistemas (1) y (2) no son equivalentes y el valor λ_0 no conviene; se haría una observación análoga para los sistemas (1) y (3).

EJERCICIOS

1. Sea $f(x) = ax^2 + bx + c$, $g(x) = a'x^2 + b'x + c'$. ¿En qué casos los sistemas

$$\begin{cases} f(x) = 0 \\ g(x) = 0 \end{cases} \qquad \begin{cases} a'f(x) - ag(x) = 0 \\ c'f(x) - cg(x) = 0 \end{cases}$$

son equivalentes? Deducir que, bajo ciertas condiciones, si estas dos ecuaciones tienen una raíz común x_0 se tiene

$$x_0 = \frac{ca' - ac'}{ab' - b\ '} = \frac{bc' - cb'}{' - ac'}.$$

2. Hallar la relación entre a, b, c para que las ecuaciones

$$x^2 - ax + 1 = 0, \qquad x^3 + bx + c = 0$$

tengan una raíz común.

3. Eliminar x entre $f(x) = 0$ y $x^2 = a$ (escribir $f(x) = f_1(x^2) + xf_2(x^2)$, siendo f_1 y f_2 dos polinomios).

213. Primeras aplicaciones de eliminación

a) Eliminación entre n ecuaciones algebraicas

Eliminar x entre las n ecuaciones algebraicas

$$f_1(x) = 0, \ \ldots, \ f_n(x) = 0$$

es hallar las condiciones verificadas por los coeficientes de estas ecuaciones, necesarias y suficientes para que estas n ecuaciones tengan al menos una raíz común.

Sea $(f, g) \to d(f, g)$ la ley interna que hace corresponder a una pareja de polinomios su m. c. d.; como en **Z** (ver § 100, ejercicio) esta ley es conmutativa y asociativa; resulta que, si se pone

$$d_1 = d(f_1, f_2) \qquad \text{y para} \qquad 2 \leq h \leq n-1 \qquad d_h = d(d_{h-1}, f_{h-1}, f_{h+1})$$

d_{n-1} es el m. c. d. de $f_1, \ldots, f_n$. Para que estos polinomios tengan una raíz común, es necesario que d_{n-1} sea de grado estrictamente positivo, para esto es necesario también lo sean $d_{n-2}, \ldots, d_1$. Recíprocamente si $d_1, \ldots, d_{n-1}$ son de grado estrictamente positivo, está claro que los polinomios $f_1, \ldots, f_n$ tienen una raíz común: habrá que escribir, pues, en general $n-1$ condiciones. En particular, si habiendo escrito que d_1 es de grado estrictamente positivo, se encuentra grd $d_1 = 1$, será suficiente escribir seguidamente los $n-2$ restantes polinomios que tienen como raíz la raíz de d_1.

Naturalmente por combinaciones de ecuaciones, y haciendo las mismas reservas que en el párrafo anterior, se puede reemplazar el sistema $f_1(x) = \ldots = f_n(x) = 0$ por un sistema más simple.

EJEMPLOS Y EJERCICIOS

1. Busquemos las condiciones para que $f(x) = 0$, de grado n, tenga una raíz múltiple de orden k $(2 \leq k \leq n)$: hay que eliminar x entre las ecuaciones

$$(1) \qquad f(x) = f'(x) = \ldots = f^{(k-1)}(x) = 0$$

y escribir seguidamente que una raíz α común a estas ecuaciones es tal que $f^{(k)}(\alpha) \neq 0$ (ver § 192, teorema 15).

Se observará que si t es una «variable de homogeneidad» (ver § 186, *e*, y § 197, ej. 2) el sistema (1) es equivalente al sistema (2) $(0 \leq h < k)$

$$(2) \qquad f^{(k-1)}_{x^{k-1}}(x) = \ldots = f^{(k-1)}_{x^{k-h-1}t^h}(x) = \ldots = f^{(k-1)}_{t^{k-1}}(x) = 0.$$

2. Escribir la condición para que $x^3 + px + q = 0$ tenga una raíz doble (ver § 192, ej. 6).

3. Escribir las condiciones para que la ecuación

$$f(x) = x^4 + ax^3 + 2x + b = 0$$

tenga una raíz triple. Utilizar los sistemas (1) y (2) (ver ej. 1 más arriba) y comparar los cálculos obtenidos.

b) Nociones sobre los sistemas de ecuaciones algebraicas con dos incógnitas

Sean f y g dos polinomios de $\mathbf{C}[X, Y]$ y consideremos el sistema

(1) $$f(x, y) = 0 \qquad g(x, y) = 0.$$

Se llama *solución del sistema* todo cero común, $(\alpha, \beta) \in \mathbf{C}$ (ver § 198), de los polinomios f y g; se tendrá, pues,

(1′) $$f(\alpha, \beta) = 0 \qquad g(\alpha, \beta) = 0.$$

Observemos primero que si f y g no son extraños, es decir, si existe un polinomio d de grado estrictamente positivo tal que $f = f_1 d$, $g = g_1 d$, todo cero (α, β) de d es una solución del sistema (1).

Supongamos f y g extraños: en consecuencia, no admitirán como divisores comunes más que elementos de $\mathbf{C}$ (ver § 195).

Si (α, β) es una solución de (1), las ecuaciones

(2) $$f(x, \beta) = 0 \qquad g(x, \beta) = 0$$

tienen al menos una solución común. Consideremos *a priori* la resultante de las ecuaciones (1) consideradas como ecuaciones en x: este resultado será un polinomio relativo a los coeficientes de estas dos ecuaciones en x (§ 211, *c*), luego un polinomio en y: $R(y)$. Las ecuaciones (2), es decir, las ecuaciones (1) para $y = \beta$, tienen por hipótesis la raíz $x = \alpha$ en común, luego $R(\beta) = 0$.

Recíprocamente si β es una raíz de $R(y) = 0$ las ecuaciones (2) tienen, al menos, una raíz común α y (α, β) es una solución de (1). Luego: *Para resolver el sistema* (1) *se forma el resultante* $R(y)$ *de los polinomios f y g considerados como polinomios en x: a toda raíz β de* $R(y) = 0$, *corresponde al menos un número α tal que (α, β) es una solución del sistema* (1).

OBSERVACIONES

1. Si f y g son distintos y de grados respectivos n y p se demuestra que el sistema (1) tiene, al menos, np soluciones; se demuestra igualmente, teniendo en cuenta «el orden de multiplicidad» de las soluciones y «las soluciones infinitas» que el sistema (1) tiene exactamente np soluciones (teorema de Bezout). La demostración de estos resultados sobrepasa el nivel de este libro.

2. Puede ocurrir que el resultante $R(y)$ sea el polinomio nulo: en cada caso donde esto ocurra se observará que f y g no son extraños.

EJEMPLOS Y EJERCICIOS

4. Discutir según los valores de los números complejos a y b el sistema (en $\mathbf{C}^2$)
$$x^2 - y^2 = 1, \qquad ax + by + c = 0.$$

5. Calcular la resultante $R(y)$ del sistema
$$x^2 + xy + 2x + 2y + a - 1 = 0, \qquad x^2 + axy + x + ay + a - 1 = 0.$$
Verificar que $R(y)$ tiene $a - 1$ como factor. Estudiar el sistema en $\mathbf{C}^2$.

6. Estudiar en $\mathbf{C}^2$ el sistema, $x^2 + y^2 - 2ax = 0$, $x^3 + y^3 - 3xy = 0$ (se podrá poner $y = tx$).

7. ¿Cómo se escribirá que las tres ecuaciones
$$f_1(x, y) = 0, \qquad f_2(x, y) = 0, \qquad f_3(x, y) = 0$$
tienen una solución común en $\mathbf{C}^2$? (f_1, f_2, f_3 elementos de $\mathbf{C}[X, Y]$).

214. Discriminante de una ecuación algebraica

Según lo que hemos visto en el § 192 (teorema 15), la ecuación $f(x) = 0$ tendrá al menos una raíz múltiple de orden $k \geq 2$ si las ecuaciones

$$f(x) = 0, \qquad f'(x) = 0$$

tienen al menos una raíz común.

Luego obtendremos la condición necesaria y suficiente para que $f(x) = 0$ tenga al menos una raíz múltiple anulando la resultante de f y f'

$$f(x) = a_0(x - \alpha_1) \dots (x - \alpha_n) \qquad (a_0 \neq 0)$$

$$f'(x) = a_0 \sum_{i=1}^{n} (x - \alpha_1) \dots (x - \alpha_{i-1})(x - \alpha_{i+1}) \dots (x - \alpha_n)$$

de donde

$$f'(\alpha_i) = a_0(\alpha_i - \alpha_1) \dots (\alpha_i - \alpha_{i-1})(\alpha_i - \alpha_{i+1}) \dots (\alpha_i - \alpha_n)$$

Según la fórmula (1) del § 211, *c*, tendremos

$$\mathrm{R} = (a_0)^{n-1} f'(\alpha_1) \dots f'(\alpha_n) = (a_0)^{2n-1} \prod_{i \neq j} (\alpha_i - \alpha_j)$$

en el producto $\prod_{i \neq j} (\alpha_i - \alpha_j)$, hay $n(n-1)$ factores, cada uno de ellos se halla escrito dos veces (en la forma $\alpha_i - \alpha_j$ y en la forma $\alpha_j - \alpha_i$), luego

$$\mathrm{R} = (a_0)^{2n-1}(-1)^{\frac{n(n-1)}{2}} \prod_{i<j} (\alpha_i - \alpha_j)^2.$$

Teniendo en cuenta el hecho de que R = S (S, determinante de SYLVESTER, § 211, ej. 8) se ve que el polinomio R = S tiene a_0 como factores (aquí $b_0 = na_0$, considérese la primera columna de S), se pone $\mathrm{R} = a_0\mathrm{D}$, el polinomio D (relativo a los coeficientes de f) se llama el *discriminante* de f y se tiene

$$\mathrm{D} = (a_0)^{2n-2}(-1)^{\frac{n(n-1)}{2}} \prod_{i<j} (\alpha_i - \alpha_j)^2$$

siendo D nulo, $f(x) = 0$ y $f'(x) = 0$ tienen al menos una raíz común, es decir, su m. c. d. Δ es de grado estrictamente positivo. Una raíz α de $\Delta(x) = 0$ es una raíz múltiple de $f(x) = 0$; su orden es el menor entero k tal que $f^{(k)}(\alpha) \neq 0$ (ver § 192, teorema 15).

EJERCICIOS

1. Calcular el discriminante de $f(x) = ax^2 + bx + c$ (se observará que con nuestras anotaciones $\mathrm{D} = 4ac - b^2$).
2. Calcular el discriminante de $f(x) = x^3 + px + q$ (se halla $\mathrm{D} = 4p^3 + 27q^2$, ver § 192, ej. 6; § 210, ej. 2, y § 213, ej. 2).

215. Transformación de ecuaciones algebraicas

Siendo f un polinomio de C[X] de grado n, cuyas raíces son $\alpha_1, \ldots, \alpha_n$ y r una fracción racional de C(X), para la cual $\alpha_1, \ldots, \alpha_n$ son sustituibles, se llama *transformada de la ecuación $f(x) = 0$ por la fracción racional r, la ecuación algebraica que tiene por raíces $r(\alpha_1), \ldots, r(\alpha_n)$.*

Sea (1) $f(x) = a_0(x - \alpha_1) \ldots (x - \alpha_n)$ $(a_0 \neq 0)$.

Pongamos

$$y = r(x), \qquad \beta_i = r(\alpha_i) \qquad (i = 1, \ldots, n)$$

la transformada de $f(x) = 0$ por r será

$$(2) \qquad g(y) = (y - \beta_1) \ldots (y - \beta_n) = 0.$$

El polinomio g podría obtenerse calculando los números $\tau_1, \ldots, \tau_n$

$$(i = 1, \ldots, n) \qquad \tau_i = \Sigma\beta_1\beta_2 \ldots \beta_i = \Sigma r(\alpha_1)r(\alpha_2) \ldots r(\alpha_i)$$

que son las funciones simétricas racionales de las raíces de la ecuación $f(x) = 0$: el cálculo, en general, es pesado; por el contrario, el cálculo del polinomio g, que vamos a efectuar, nos dará los valores de las funciones simétricas $\tau_1, \ldots, \tau_n$ (ver ej. 3 más abajo).

Pongamos $r = u/v$, siendo u/v irreducible, β será una raíz de (2) si y sólo si existe un valor α tal que

$$f(\alpha) = 0, \qquad \beta v(\alpha) - u(\alpha) = 0.$$

Luego se obtendrá la transformada de $f(x) = 0$ por r eliminando x entre las dos ecuaciones

$$f(x) = 0, \qquad yv(x) - u(x) = 0$$

es decir, $g(y)$ es el resultante de los dos polinomios en x: $f(x)$ e $yv(x) - u(x)$.

Observemos que dos fracciones distintas r y s pueden dar la misma transformada para la ecuación $f(x) = 0$: es suficiente para esto que se tenga

$$(i = 1, \ldots, n) \qquad s(\alpha_i) = r(\alpha_i)$$

esta observación puede tener un gran interés práctico (ver ej. 2 más abajo).

La transformación de las ecuaciones es principalmente útil para la *resolución de las ecuaciones.* Si $f(x) = 0$ es de grado n, $g(y) = 0$ es también de grado n. Pero ciertas propiedades de $f(x) = 0$ pueden llevar consigo la elección de r de manera que $g(y) = 0$ tenga raíces múltiples; en este caso la ecuación $g(y) = 0$ de grado estrictamente inferior a n llamada "*resolvente*" de $f(x) = 0$ para la transformación $y = r(x)$ (ver ej. 4 y 5 más abajo y ej. 425 y 426 final del capítulo).

EJEMPLOS Y EJERCICIOS

1. La transformada de $f(x) = 0$ por $y = (ax + b)/(cx + d)$ es

$$g(y) = (a - cy)^n f\left(\frac{dy - b}{a - cy}\right)$$

en particular:



a) $y = x + h$ (traslación de las raíces) da $g(y) = f(y - h)$, ecuación que se puede obtener mediante la fórmula de TAYLOR.

b) $y = kx$ ($k \neq 0$, homotecia de las raíces) da $g(y) = f(y/k)$. En particular la ecuación $f(-x) = 0$ se llama *«ecuación en las opuestas»* de las raíces de $f(x) = 0$.

c) $y = 1/x$ para $f(x) = a_0x^n + \dots + a_n = 0$ ($a_0a_n \neq 0$) da $g(x) = a_nx^n + \dots + a_0 = 0$, llamada *«ecuación en las inversas»* de las raíces de $f(x) = 0$.

2. *a*) Demostrar que para obtener la transformada de $f(x) = 0$ para $y = u(x)/v(x)$ se puede reemplazar u y v por dos polinomios u_1 y v_1 de grado estrictamente inferior al de f (u_1 y v_1 son los restos respectivos en las divisiones euclídeas de u y v por f).

b) Demostrar que se puede reemplazar $y = u(x)/v(x)$ por $y = w(x)$, siendo w un polinomio de grado estrictamente inferior al de f (observar que v y f son extraños).

3. Si α_1, α_2, α_3 son las raíces de $f(x) = x^3 + px + q = 0$ ($p + q + 1 \neq 0$) se pone $\beta_i = (\alpha_i + 1)/(\alpha_i - 1)$, calcular en función de p y q,

$$\beta_1 + \beta_2 + \beta_3, \qquad \beta_2\beta_3 + \beta_3\beta_1 + \beta_1\beta_2, \qquad \beta_1\beta_2\,\beta_3$$

(transformar $f(x) = 0$ por $y = (x + 1)/(x - 1)$).

4. *Ecuaciones bicuadradas generalizadas.* Sea $f(x) = 0$ una ecuación tal que $f(\alpha) = 0$ implica $f(-\alpha) = 0$; se supone, además, $f(0) \neq 0$.

a) Demostrar que f es de grado $n = 2m$ y que $f(x) = h(x^2)$. Deducir de ello que la resolución de $f(x) = 0$ se reduce a la de $h(y) = 0$, que es de grado $m = n/2$ y al cálculo de m raíces cuadradas.

b) $h(y) = 0$, llamada *«resolvente de* $f(x) = 0$*»*, no es la transformada de $f(x) = 0$ por $y = x^2$. Formar esta transformada que es de grado $n = 2m$. ¿Se podría prever el resultado obtenido? (utilizar § 212, ej. 3).

5. *Ecuaciones recíprocas.* Sea $f(x) = 0$ una ecuación tal que $f(0) \neq 0$ y tal que $f(\alpha) = 0$ implica $f(1/\alpha) = 0$. Se supone, además, que $f(1)f(-1) \neq 0$.

a) Demostrar que f es de grado $n = 2m$ y que para $k = 0, \dots, n$ se tiene $a_k = a_{n-k}$.

b) Demostrar que la transformada de $f(x) = 0$ por $y = x + 1/x$, sea $g(y) = 0$, sólo tiene raíces dobles. Deducir la existencia de una *«resolvente»* $h(y) = 0$ de grado m (se observará que si se pone para $k \geq 0$, $y_k = x^k + 1/x^k$, se tiene para $k \geq 2$: $y_k = y_{k-1}\, y_1 - y_{k-2}$).

Ejercicios

N. B. — *Salvo mención contraria se supone que todas las ecuaciones consideradas son con coeficientes complejos.*

395. ¿Qué relaciones deben verificar los coeficientes reales de la ecuación $a_0x^n + ... + a_n = 0$ para que sus raíces α_h sean tales que

$$(h = 0, \ldots, n) \qquad \alpha_h = \operatorname{tg}\left(\varphi + \frac{h\pi}{n}\right) \qquad (\varphi \in \mathbf{R}).$$

Terminar los cálculos para $n = 3$.

396. Determinar y resolver las ecuaciones de tercer grado cuyas raíces están en progresión geométrica.

397. Determinar a para que dos de las raíces de la ecuación

$$3x^4 + ax^3 + 2x^2 + 12x - 8 = 0$$

tenga un producto igual a 4. Resolver entonces la ecuación (utilizar § 209, ej. 3).

398. Determinar a para que las cuatro raíces α_1, α_2, α_3, α_4 de

$$x^4 - 6x^3 + 8x^2 + ax + 25 = 0,$$

verifiquen $\alpha_1 + \alpha_2 = \alpha_3 + \alpha_4$. Resolver entonces la ecuación (utilizar § 209, ej. 3).

399. Determinar a para que la ecuación

$$(a + 3)x^3 - ax^2 - (a + 2)x + a = 0$$

tenga una raíz de módulo 1.

400. Formar las ecuaciones de tercer grado cuyas raíces son los afijos de los vertices:

a) de un triángulo isósceles,
b) de un triángulo rectángulo,
c) de un triángulo rectángulo isósceles.

401. Formar las ecuaciones de cuarto grado cuyos afijos sean los vértices de un rectángulo.

402. Demostrar que las raíces α_1, α_2, α_3, α_4 de la ecuación

$$a_0x^4 + 4a_1x^3 + 6a_2x^2 + 4a_3x + a_4 = 0$$

son tales que $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = -1$ (V. § 124, *d*) si y sólo si

$$\begin{vmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{vmatrix} = 0.$$

(Se observará que $(\alpha_1 + \alpha_2)\ 2(\alpha_1\alpha_2 + (\alpha_3 + \alpha_4) = \alpha_3\alpha_4)$ y se utilizará ej. 3, § 209).

403. Demostrar que la condición encontrada en el ejercicio precedente es la misma para la cual existe α, β, a, $b \in \mathbf{C}$, tales que

$$a_0X^4 + 4a_1X^3 + 6a_2X^2 + 4a_3X + a_4 = a(X - \alpha)^4 + b(X - \beta)^4.$$

(Observar que la función homográfica conserva la razón doble de cuatro números, V. § 124, *d*).

404. Siendo α, β, γ las raíces de $x^3 + px + q = 0$ $(q \neq 0)$ demostrar que cuando se permuta de todas las maneras posibles las raíces $\frac{\beta}{\gamma} + \frac{\gamma}{\alpha} + \frac{\alpha}{\beta}$ sólo toman dos valores a y b. Calcular $a + b$ y ab en función de p y q.

405. Si α, β, γ son las raíces de $x^3 + px^2 + qx + r = 0$, calcular en función de p, q, r, cuando están definidas, las expresiones

$$\text{a) } \sum \frac{\alpha + 1}{(\beta + 2)(\gamma + 2)}, \qquad \text{b) } \sum \frac{\alpha^2}{\beta^2}, \qquad \text{c) } \sum \frac{1}{(\alpha - \beta)^2}.$$

406. Siendo (α_h) $(1 \leq h \leq 6)$ las raíes de $x^6 - x - 1 = 0$. Calcular

$$\sum \frac{(\alpha_1)^2 + 1}{(\alpha_1 - 1)(\alpha_1 - 2)}.$$

407. Si α, β, γ, δ son las raíces de $x^4 - 3x^2 + x - 1 = 0$, calcular $\sum \alpha^3\beta^2\gamma^2$.

a) Aplicando el método general.

b) Observando, previamente, que $\alpha^3\beta^2\gamma^2 = \alpha^4\beta^2\gamma^2\delta^2(\alpha/\delta^2)$.

408*. *Demostración del teorema de* D'ALEMBERT-GAUSS (§ 193, **b**).
Sea $f \in \mathbf{C}[X]$, grd $f = n \geq 1$. Se suponen admitidos los resultados siguientes:

1. Todo número real > 0 tiene dos raíces cuadradas reales;
2. Todo polinomio de $\mathbf{R}[X]$ y de grado impar tiene al menos una raíz real (ver curso de Análisis);
3. Para todo polinomio f de $\mathbf{C}[X]$ existe un subcuerpo conmutativo Δ de $\mathbf{C}$ en el que f tiene todas sus raíces $\alpha_1, \alpha_2, \ldots, \alpha_n$ (V. ej. 348, cap. 11).

a) Demostrar que el hecho que todo polinomio de $\mathbf{R}[X]$ tenga una raíz en $\mathbf{C}$ implica que debe suceder lo mismo para todo polinomio de $\mathbf{C}[X]$ (si $f \in \mathbf{C}[X]$, observar que $f\bar{f} \in \mathbf{R}[X]$);

b) Sea $f \in \mathbf{R}[X]$, grd $f = n > 1$. Colocándose en Δ se pone

$$\beta_{hk} = \alpha_h + \alpha_k + a(\alpha_h\alpha_k) \ (h < k,\ a \in \mathbf{R}).$$

Demostrar que el polinomio f_a que tiene por raíces $(\beta_{hk})(1 \leq h < k \leq n)$ pertenece $\mathbf{R}[X]$ y es de grado $n(n-1)/2$.

c) Demostrar que todo entero natural $n \geq 1$ puede escribirse $n = 2^pq$, p y q enteros naturales, q *impar* y que

$$\frac{n(n-1)}{2} = 2^{p-1}q' \qquad (q' \text{ impar}).$$

d) Sea $f \in \mathbf{R}[X]$ grd $n = 2^pq$ (q impar), demostrar por recurrencia sobre p que f tiene al menos una raíz en $\mathbf{C}$.

(El resultado es verdadero para $p = 0$; hipótesis de recurrencia; es verdadero para $p - 1$ luego f_a tiene una raíz en $\mathbf{C}$ para todo a. Dando a a estrictamente más de $n(n-1)/2$ valores y utilizando el principio de los «cajones» (§ 35, ejercicio) se demostrará que existe a y a' números reales y distintos y un par (h, k) de enteros naturales tales que

$$\beta_{hk} = \alpha_h + \alpha_k + a\alpha_h\alpha_k \in \mathbf{C}, \qquad \beta_{hk} = \alpha_h + \alpha_k + a'\alpha_h\alpha_k \in \mathbf{C}).$$

e) Concluir para $f \in \mathbf{C}[X]$.

409*. Siendo **K** un cuerpo conmutativo completamente ordenado, demostrar que las propiedades siguientes son equivalentes:

a) $\mathbf{K}(i)$ $(i^2 + 1 = 0)$ es algebraicamente cerrado.

b) Todo elemento de **K** es cuadrado de un elemento **K** y todo polinomio de grado impar de $\mathbf{K}[X]$ tiene una raíz en **K**.

(Utilizar capítulo 6, ej. 122; cap. 11, ej. 348 y el ejercicio precedente.)

410. Sea f un polinomio de grado n de $\mathbf{C}[X]$, se designa por f_k, el polinomio que tiene por raíces simples todas las raíces de orden h de f, es decir

$$f = f_1(f_2)^2 \dots (f_p)^p \qquad n = \sum_{k=1}^{p} k \text{ grd } f_k,$$

Si f no tiene raíces de orden k se pondrá $f_k = 1$.
Se designa por $g_0, g_1, \dots, g_p$ los polinomios definidos de la manera siguiente: $g_0 = f$ y para $1 \leq k \leq p$, g_k es el m.c.d. de g_{k-1} y de g'_{k-1}.
Se pone finalmente: $h_h = g_{k-1}/g_k$ $(1 \leq k \leq p)$.

a) Calcular $g_1, g_2, \dots, g_p$;

b) Calcular $h_1, h_2, \dots, h_p$, deducir de ello $f_1, f_2, \dots, f_p$.

411. Calcular los polinomios f_k (V. ej. anterior) para los polinomios

a) $X^5 - 8X^3 + 16X^2 + 36X - 32$

b) $X^7 + X^6 + 2X^4 + 2X^3 + X + 1$

412. Determinar las ecuaciones

$$x^4 + ax^3 + bx^2 + cx + d = 0$$

que tienen dos raíces dobles.

413. Determinar a, b, c, para que la ecuación

$$x^5 + ax^4 + bx^2 + c = 0$$

tenga una raíz triple.

414. Determinar a y b para que la ecuación

$$x^4 + 4x^3 - 2x^2 + ax + b = 0$$

sólo tenga dos raíces distintas.

415. Si A, B, C son dos polinomios de $\mathbf{C}[X]$, demostrar que todo polinomio P de $\mathbf{C}[X]$ verificando $AP'' + BP' + PC = 0$ no puede tener raíces múltiples, salvo eventualmente las raíces de A.

416. Eliminar x entre las dos ecuaciones.

$$4x^5 - 25x^3 + 10x^2 + 21x - 10 = 0, \qquad 2x^3 + 5x^2 + x - 2 = 0.$$

417. Siendo u, v, w tres polinomios de $\mathbf{R}[X]$ primos en su conjunto y $t \in T$, con T el conjunto de los números reales menos las raíces de w se considera la aplicación φ de T en $\mathbf{R}^2$ definido por

$$(1) \qquad x = \frac{u(t)}{w(t)}, \qquad y = \frac{v(t)}{w(t)}.$$

Si t describe T, (x, y) describe una curva Γ de la que las ecuaciones (1) constituyen una representación paramétrica.

a) Demostrar que la ecuación cartesiana de Γ, es decir, la condición necesaria y suficiente para que $(x, y) \in \Gamma$ se obtiene eliminando t entre las ecuaciones

(2) $$u(t) - xw(t) = 0, \qquad v(t) - yw(t) = 0.$$

b) Demostrar que los puntos múltiples de (Γ) se obtienen al escribir que las ecuaciones (2) tienen al menos dos soluciones comunes en t.

c) Hallar las ecuaciones cartesianas y los puntos múltiples de dos curvas

α) $$x = \frac{2t-1}{t^2-1}, \qquad y = \frac{t^2}{t-1}.$$

β) $$x = \frac{t+2}{t(t^2-1)}, \qquad y = \frac{t}{t^2-1}.$$

Resolver los sistemas

418. $\begin{cases} x^3 - 3xy^2 = 1 \\ 3x^2y - y^3 = \sqrt{3}. \end{cases}$ **419.** $\begin{cases} x^3 - 2xy^2 - y^3 - 1 = 0 \\ x^2 + 3xy + y^2 + 1 = 0. \end{cases}$

420. Formar la ecuación $g(y) = 0$ que tienen por raíces los cuadrados de las raíces de la ecuación $f(x) = 0$:

a) Eliminando x entre $f(x) = f_0(x^2) + xf_1(x^2) = 0$ y $x^2 - y = 0$.

b) Demostrando que $g(y) = 0$ es equivalente a la ecuación que se obtiene reemplazando x^2 por y en $f(x)f(-x) = 0$.

421. Formar la ecuación que tiene por raíces las potencias m-ésimas de las raíces de $f(x) = 0$. (Escribir $f(x) = f_0(x^m) + xf_1(x^m) + \ldots + x^{m-1}f_{m-1}(x^m)$, V. ej. 322, cap. 11 y observar que

$$f(\omega_0 x)f(\omega_1 x) \ldots f(\omega_{m-1}x) = g(x^m),$$

siendo g un polinomio y ω_h $(0 \leq h \leq m-1)$ las raíces m-ésimas de 1.)

422 Si $R(u, v)$ es una fracción racional para la cual toda pareja (α_h, α_k) de raíces de la ecuación $f(x) = 0$ es sustituible en R, demostrar que la ecuación que tiene por raíces $R(\alpha_h, \alpha_k)$ $(h \neq q)$ se obtiene eliminando u y v entre

$$f(u) = 0, \qquad f(v) = 0, \qquad R(u, v) - y = 0.$$

Efectuar los cálculos para $f(x) = x^3 + px + q$ y $R = u - v$.

423. Si $f(x) = 0$ tiene por raíces $\alpha_1, \alpha_2, \ldots, \alpha_n$ se propone formar las ecuaciones que tienen por raíces:

a) las diferencias de las raíces (α_h) dos a dos; **b)** las sumas de las raíces α_h, dos a dos.

Demostrar que estas ecuaciones se obtienen eliminando, respectivamente, y y g entre las ecuaciones

$$f\left(\frac{y+z}{2}\right) = 0, \qquad f\left(\frac{z-y}{2}\right) = 0.$$

Efectuar todos los cálculos para $f(x) = x^3 + px + q$.

424. Sea $f(x) = x^3 + px + q = 0$ y $R(u, v)$ una fracción racional simétrica, demostrar que el hallar la ecuación que tiene por raíces $R(\alpha_h, \alpha_k)$(donde α_1, α_2 son las raíces de $f(x) = 0$ y $h \neq k$) se reduce a encontrar la transformada de $f(x) = 0$ por $y = R_1(x)$, siendo R_1 una fracción racional. (Observar que existe una fracción racional S tal que $R(u, v) = S(u + v, uv)$.)

Formar las ecuaciones que tienen por raíces:

a) $\alpha_2 + \alpha_3$, $\alpha_3 + \alpha_1$, $\alpha_1 + \alpha_2$ (compararlo a un resultado del ejercicio 423).

b) $(\alpha_2 - \alpha_3)^2$, $(\alpha_3 - \alpha_1)^2$, $(\alpha_1 - \alpha_2)^2$, ¿se podía prever **la** forma del término constante de la ecuación obtenida? (V. § 210, ej. 2 y § 214, ej. 2.)

425*. Se considera la ecuación

(E) $$x^8 - 8x^6 + 20x^4 - 16x^2 - x + 2 = 0$$

y se pone

$$R(x) = x^2 - 2, \qquad R_0(x) = x \quad \text{y} \quad R_n(x) = R[R_{n-1}(x)] \quad \text{para} \quad n \geq 1.$$

1.° Demostrar que para todo $n \geq 1$ las ecuaciones $R_n(x) = R_0(x)$ tienen dos raíces comunes independientes de n, α y β.

2.° **a)** Formar la transformada de (E) para $y = R(x)$. Demostrar que (E) admite α y β por raíces. Se designa por (F) la ecuación que tiene por raíces las raíces de E distintas de α y β. Demostrar que (F) es de grado seis.

b) Demostrar que las raíces de (F) se dividen en dos ciclos de tres valores, las raíces de cada ciclo se permutan por la transformación $y = R(x)$.

c) Siendo a una raíz de (F), demostrar que $R_0(a) + R_1(a) + R_2(a)$ sólo toma dos valores. Deducir que (F) se descompone en dos ecuaciones (F_1) y (F_2) de grado tres que se formará.

3.° Se pone $2 \cos \theta = x$ y $f(x) = 2 \cos 8\theta$, demostrar que la ecuación $f(x) - x = 0$ es equivalente a (E). Hallar de nuevo los resultados de la pregunta segunda.

426*. Se considera la función homográfica propia con puntos invariantes α y β distintos (V. § 124).

$$y = R(x) = -\frac{ax + c}{x + b}$$

y se llama ecuación E toda ecuación $f(x) = 0$ de grado n, invariante para $y = R(x)$ tal que $f(\alpha)f(\beta) \neq 0$.

a) Escribir la relación homográfica en la forma

$$\frac{y - \alpha}{y - \beta} = k\frac{x - \alpha}{x - \beta}.$$

calcular $k + 1/k$ en función de a, b, c (V. § 124, ej. 10).

b) Demostrar que si existen ecuaciones E, a, b, c verifican una condición que se supondrá en todo lo que sigue.

c) Demostrar que

$$(x + a)^n f\left(-\frac{bx + c}{x + a}\right) = sf(x)$$

calcular s.

d) Demostrar que la transformación $z = (x - \alpha)/(x - \beta)$ reduce, en general, la resolución de E a la resolución de una ecuación de menor grado y a la extracción de un cierto número de raíces m-ésimas.

Estudiar el caso de n primo.

e) Hallar todas las ecuaciones de grado 4 invariantes por $y = x/(x-1)$. Resolver $x^4 - 3x^3 + x^2 + 4x - 2 = 0$.

f) Hallar todas las ecuaciones de grado 3 invariantes por $y = 1/(x-1)$.

427. Sea $f(x) = 0$ una ecuación con coeficientes reales, demostrar utilizando el teorema de ROLLE (ver curso de Análisis) y la descomposición de f en $\mathbf{R}[X]$ (V. § 193, c) los resultados siguientes:

a) Entre dos raíces reales consecutivas de $f(x) = 0$, hay un número impar de raíces reales de $f'(x) = 0$.

b) Entre dos raíces reales consecutivas de $f'(x) = 0$, existe a lo sumo una raíz real de la ecuación $f(x) = 0$.
$f(x) = 0$ tiene a lo sumo una raíz real menor (resp. mayor) que la menor (resp. mayor) raíz real de $f'(x) = 0$.

c) Si $\alpha < \beta$ son dos raíces reales consecutivas de la ecuación $f'(x) = 0$:
Si $f(\alpha)f(\beta) < 0$ hay una y sola una raíz real de $f(x) = 0$ en $]\alpha, \beta[$.
Si $f(\alpha)f(\beta) > 0$ no hay ninguna raíz real de $f(x) = 0$ en $]\alpha, \beta[$.

428. Sea $f(x) = a_0x^n + \ldots + a_n = 0$ una ecuación con coeficientes reales, se designan por $\alpha_1 < \alpha_2 < \ldots < \alpha_m$ las raíces reales de $f'(x) = 0$ y se llama «sucesión de ROLLE» de $f'(x) = 0$ la sucesión de números reales

$$(-1)^n a_0 \qquad f(\alpha_1), \qquad f(\alpha_2), \ \ldots, \ f(\alpha_m), \qquad a_0.$$

Demostrar que todas las raíces de $f(x) = 0$ son reales y distintas si y sólo si $f'(x) = 0$ tiene sus $n - 1$ raíces reales y distintas y si la sucesión de ROLLE de $f(x)$ presenta n cambios de signo.
Aplicar este resultado a $x^3 + px + q = 0$.

429. Demostrar que la ecuación

$$1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} = 0$$

tiene una sola raíz real para n impar y ninguna para n par. (Razonar por inducción utilizando el ejercicio 427.)

430. Se designa por P_n y Q_n polinomios con coeficientes reales verificando las condiciones

$$P_n = 2XP_{n-1} - P_{n-2}, \qquad (P_0 = 1, \qquad P_1 = 2X)$$
$$Q_n = 2XQ_{n-1} - Q_{n-2}, \qquad (Q_0 = a, \qquad Q_1 = bX + c)$$

$(a, b, c \in \mathbf{R})$.

a) Demostrar que $P_n(x) = 0$ tiene todas sus raíces reales distintas y separadas por las de $P_{n-1}(x) = 0$.

b) En qué se transforma $P_n(x)$ cuando se hace el cambio de variable $x = \cos\theta$. Hallar de nuevo así el resultado de la pregunta anterior.

c) Demostrar que Q_n se expresa linealmente en función de P_n, P_{n-1}, P_{n-2}. Cómo se debe escoger a, b, c para que $Q_n(x) = 0$ tenga todas sus raíces reales y separadas por las de $Q_{n-1}(x) = 0$. (Ingreso en l'École Polytechnique).

431. Se consideran los polinomios $f_n(X) = D^n(X^2 - 1)^n$ (donde D^n es el operador de derivación n-ésima).

a) Demostrar que $g_n(X) = (X^2 - 1)^n$ verifica

$$(X^2 - 1)g_n''(X) - 2(n-1)Xg_n'(X) - 2ng_n(X) = 0$$

deducir

$$(X^2 - 1)f_n''(X) + 2Xf_n'(X) - n(n+1)f_n(X) = 0.$$

b) Deducir que $f_n(x) = 0$ tiene n raíces reales distintas y que estas raíces pertenecen a $]-1, 1[$.

432. Sea f un polinomio con coeficientes reales de grado 3, demostrar que la ecuación

$$[f'(x)]^2 - 2f(x)f''(x) = 0$$

tiene dos raíces reales y dos raíces complejas conjugadas.

433. Siendo f un polinomio de grado n con coeficientes reales y teniendo todas sus raíces reales y distintas demostrar que la ecuación $f(x)f''(x) - [f'(x)]^2 = 0$ no tiene raíces reales. Utilizar la relación

$$\left(\frac{f'(x)}{f(x)} = \frac{1}{x - \alpha_2} + \ldots + \frac{1}{x - \alpha_n} \right).$$

434. Se designa por P_f (resp. N_f) el conjunto de las raíces estrictamente positivas (resp. negativas) de $f \in \mathbf{R}[X]$ demostrar que si para todo f se sabe hallar una cota superior de P_f, se sabrá también hallar una cota inferior > 0 de P_f, una cota inferior de N_f y una cota superior < 0 de N_f.

435. Todos los polinomios tratados son con coeficientes reales.

a) Se considera

$$f_0(x) = a_0x^n + \ldots + a_px^{n-p} - (a_{p+1}x^{p+1} + \ldots + a_n) = 0$$

con $a_h \geq 0$ $(0 \leq h \leq n)$ y $a_0a_n \neq 0$.

Demostrar que si para $\alpha > 0$ se tiene $f(\alpha) > 0$; α es una cota superior del conjunto de las raíces positivas de $f(x) = 0$ (se pondrá en $f_0(x)$, x^{n-p} en factor).

b) Escribiendo $f \in \mathbf{R}[X]$ en la forma

$$f(x) = f_1(x) + f_2(x) + \ldots + f_m(x)$$

donde los coeficientes de $f_1, f_2, \ldots, f_m$ tienen las propiedades de los coeficientes del polinomio $f_0(x)$ de la pregunta anterior, hallar una cota superior del conjunto de las raíces positivas de $f(x) = 0$.

c) Hallar las cotas superior e inferior > 0 de P_f (resp. < 0 de N_f) (V. ej. 432) para

$$f(x) = x^8 - 10x^7 + 3x^6 + x^5 - 4x^4 + x^3 + 2x^2 - 7x - 1 = 0.$$

436. Sea $f(X) = a_0X^n + a_1X^{n-1} + \ldots + a_n \in \mathbf{R}[X]$ con $a_0 > 0$.

Se designa por A el mayor valor absoluto de los coeficientes negativos.

a) Demostrar que $1 + A/a_0$ es una cota superior del conjunto de las raíces > 0 de $f(x) = 0$. Se escribirá

$$f(x) = [a_0x^n - A(x^{n-1} + x^{n-2} + \ldots + 1)] + [(a_1 + A)x^{n-1} + \ldots + a_n + A].$$

b) Sea $n - r$ el grado del término de mayor grado de f afectado de un coeficiente < 0 demostrar que $1 + \sqrt[r]{A/a_0}$ es una cota superior del conjunto de las raíces > 0 de $f(x) = 0$. Se escribirá

$$f(x) = [a_0x^n - A(x^{n-r} + x^{n-r-1} + \dots + 1)] \\ + [a_1x^{n-1} + \dots + a_{r-1}x^{n-r+1}] + [(a_r + A)x^{n-r} + \dots + a_n + A)].$$

437. Sea f un polinomio con coeficientes reales de grado n, demostrar que si para a real, $f(a)$, $f'(a)$, ..., $f^n(a)$ son reales positivos, $x > a$ implica $f(x) > 0$. Deducir de lo anterior un método para determinar una cota superior del conjunto de las raíces > 0 de $f(x) = 0$.

438. Hallar las cotas superiores del conjunto de las raíces > 0 de las ecuaciones siguientes, sirviéndose de los ejercicios 435, 436, 437.

a) $f(x) = 0$ dado en el ejercicio 435.

b) $x^6 + x^5 - 2x^4 + x - 10 = 0$.

c) $x^6 + x^5 - 2x^4 + x - 1000 = 0$.

439. Sea α una raíz de $f = a_0X^n + \dots + a_n \in \mathbf{Z}[X]$ si $a_0, \dots, a_n$ son primos en su conjunto.
a) Se supone $\alpha \in \mathbf{Z}$. Demostrar que α divide a_n. Si además $f(1) \neq 0$ [resp. $f(-1) \neq 0$] demostrar que $\alpha - 1$ divide $f(1)$ [resp. $\alpha + 1$ divide $f(-1)$].
b) Se supone $\alpha = p/q \in \mathbf{Q}$ con la fracción p/q irreducible. Demostrar que p divide a a_n y que q divide a a_0
c) Calcular las raíces enteras o racionales de las ecuaciones

$$2x^3 - 12x^2 + 13x - 15 = 0$$
$$4x^3 - 8x^2 + 15x - 3 = 0$$
$$6x^4 + 19x^3 - 7x^2 - 26x + 12 = 0$$
$$x^5 - 34x^3 + 29x^2 - 212x - 300 = 0.$$

440. Resolver la ecuación

$$2x^7 + x^6 - x^5 - 6x^4 - 5x^3 - 2x^2 + x + 1 = 0$$

sabiendo que tiene una raíz racional y dos raíces dobles.

441. *Resolución de la ecuación de tercer grado.*

a) Demostrar que un cambio de variable simple permite reducir la resolución de la ecuación general de tercer grado a la resolución de

(1) $$x^3 + px + q = 0 \qquad (p, q \in \mathbf{C})$$

que se tratará en todo el problema (V. § 209, ej. 5).

b) Demostrar que existen a, b, α, β tales que

$$X^3 + pX + p = a(X - \alpha)^3 + b(X - \beta)^3$$

deducir de ello un método de resolución de (1).

c) En (1) se pone $x = u + v$ con $3uv + p = 0$. Demostrar que u^3 y v^3 son las raíces y_1, y_2 de una ecuación de segundo grado que se formará. ¿Cómo hay que asociar las raíces cúbicas de y_1 e y_2 para que su suma sea una raíz de (1)? (Fórmulas de CARDANO.)

Efectuar todos los cálculos cuando p y q son reales (distinguir los casos $4p^3+27q^2>0$ y $4p^3+27q^2<0$). Resolver las ecuaciones

$$x^3-2x-12=0, \qquad x^3-15x-4=0.$$

d) Si α, β, γ son las raíces de (1) demostrar que $z=(\alpha+\beta j+\gamma j^2)^3$, ($j^3=1$) sólo toma dos valores z_1 y z_2 cuando se permuta α, β, γ de todas las maneras posibles. Formar la ecuación de segundo grado admitiendo z_1 y z_2 como raíces. Deducir de lo anterior un método de resolución de la ecuación (1). Efectuar todos los cálculos cuando p y q son reales. Resolver las dos ecuaciones de la pregunta **c**).

e) Se supone p y q reales y $4p^3+27q^2<0$. Demostrar que existe λ y α reales tales que las raíces de (1) sean

$$\lambda\cos\alpha \qquad \lambda\cos\left(\alpha+\frac{2\pi}{3}\right), \qquad \lambda\cos\left(\alpha+\frac{4\pi}{3}\right).$$

Calcular λ y $\cos 3\alpha$ en función de p y q.
Resolver con ayuda de una tabla $x^3-4x+1=0$.

442. *Resolución de la ecuación de cuarto grado.*

a) Demostrar que un cambio de variable simple permite convertir la resolución de la ecuación general de cuarto grado a la resolución de la

$$(1) \qquad x^4+px^2+qx+r=0 \qquad (p,\ q,\ r\in \mathbf{C})$$

que se considerará en todo el problema (V. § 209, ej. 5).

b) Se escribe

$$X^4+pX^2+qX+r=(X^2+\lambda)^2+g(X).$$

Calcular λ para que g sea el cuadrado de un binomio. Deducir que la resolución de una ecuación de cuarto grado se reduce a la resolución de una ecuación de tercer grado y de dos ecuaciones de segundo grado.
Resolver $x^4+3x^2-2x+2=0$.

c) Sea α_1, α_2, α_3, α_4 las raíces de (1), demostrar que, siendo σ una permutación cualquiera de $\{1, 2, 3, 4\}$,

$$y_\sigma=\alpha_{\sigma(1)}\alpha_{\sigma(2)}+\alpha_{\sigma(3)}\alpha_{\sigma(4)}$$

sólo toma tres valores. Formar la ecuación de tercer grado $g(y)=0$ que tiene estos tres valores como raíces. Si β es una raíz de esta ecuación, demostrar que las relaciones entre los coeficientes y las raíces de (1) junto con la relación $\alpha_1\alpha_2+\alpha_3\alpha_4=\beta$ permiten de resolver (1) (V. § 209, ej. 3). Resolver la ecuación dada en la pregunta **b**).

14

VALORES Y VECTORES PROPIOS DE UN ENDOMORFISMO REDUCCION DE MATRICES

En todo este capítulo K representa un cuerpo conmutativo, E un espacio vectorial sobre K. Para simplificar la escritura pondremos $\mathrm{id}_E = e$, indicando el contexto de que en el espacio e es la aplicación idéntica.

216. Introducción

Sea f un endomorfismo de un espacio vectorial E de dimensión n sobre el cuerpo conmutativo K. Vamos a buscar una base de E tal que la matriz de f con relación a esta base sea lo más "simple" posible. Observemos que si existe un subespacio propio F estable para f y si se toma una base (a_i) $(1 \leq i \leq n)$ tal que $(a_1, \ldots, a_k)$ $(0 < k < n)$ sea una base de F, A tomará la forma

$$A = M(f, (a_i)) = \begin{pmatrix} A'A'' \\ O\,A''' \end{pmatrix}$$

en donde el bloque $A' = M(g, (a_i))$, siendo g la aplicación lineal[37] inducida por f en F.

De lo que resulta que si existe una familia de subespacios (F_j) $(1 \leq j \leq m)$ estables por f tal que

$$E = F_1 \oplus \ldots \oplus F_m$$

y si se toma por base (a_i) $(1 \leq i \leq n)$ la reunión de las bases de $F_1, \ldots, F_m$ en el orden: base de F_1, base de F_2, ..., base de F_m, A toma la forma

$$A = M(f, (a_i)) = \begin{pmatrix} {}_1 & O & \ldots & O \\ O & A_2 & \ldots & O \\ \vdots & \vdots & & \vdots \\ O & O & \ldots & A_m \end{pmatrix}.$$

Es decir, A es una matriz, *tabla cuadrado diagonal de matrices cuadradas.*

(37) Se recuerda que si $f(E) \subset F$, la aplicación g de F en sí mismo que coincide con f sobre F se llama *aplicación inducida* por f sobre F. Algunos autores la llaman la restricción de f en F; es un abuso de lenguaje, pues g aplica F en F y la restricción f_F aplica F en E.

Estamos, pues, obligados a "simplificar" las matrices A_j de orden estrictamente inferior al de $M(f)$.

Las matrices escalares (que comprenden la matriz unidad y la matriz cuadrada cero) se consideran como las más "simples" de todas las matrices de su orden, nos vemos encaminados a buscar si no se podría encontrar subespacios F_j estables para f, cuya suma directa sea E y tales que la aplicación g_j, inducida sobre cada uno de ellos por f sea una homotecia, es decir, tales que

$$x \in F_j \Rightarrow f(x) = \lambda_j x \qquad (\lambda_j \in K)$$

en este caso, en efecto, si $h_j = \dim F_j$

$$A_j = M(g_j) = \lambda_j F_{h_j}$$

y finalmente A, matriz de f relativamente a la base, reunión de las bases de los F_j, sería un cuadro diagonal de matrices escalares, es decir, una *matriz diagonal*. Vamos a ver que no es siempre posible y daremos las condiciones para que sea así: se dice entonces que f es *diagonalizable*. En el caso general indicaremos en parte en la teoría, en parte en los ejercicios al final del capítulo, cómo se puede encontrar unas bases tales que, respecto a ellas, $M(f)$ sea relativamente "simple".

Sin embargo, prestaremos más interés al caso de los vectores $x \neq 0$ tales que haya un λ de K que verifique $f(x) = \lambda x$.

217. Vectores propios y valores propios de un endomorfismo

a) DEFINICIÓN 1. — *Siendo f un endomorfismo de un espacio vectorial* E *sobre* K, *se llama* vector propio *de f todo vector x tal que existe un elemento λ de* K *verificando*

$$(1) \qquad f(x) = \lambda x.$$

Si $x = 0$, la relación (1) se verifica cualquiera que sea λ; supongamos $x \neq 0$ verificando (1), el escalar λ es entonces único, pues

$$(x \neq 0 \quad \text{y} \quad \lambda x = \lambda' x) \Rightarrow (\lambda - \lambda')x = 0 \Rightarrow \lambda - \lambda' = 0$$

de donde la definición:

DEFINICIÓN 2. — *Si f es un endomorfismo de un espacio vectorial* E *sobre* K, *se llama* valor propio *de f todo elemento λ de* K *tal que existe un vector $x \neq 0$ de* E *que verifica*

$$(1) \qquad f(x) = \lambda x$$

Si λ es un valor propio, por definición el conjunto de los vectores x de E que verifican (1) sea $V(\lambda)$ es distinto de $\{0\}$. Sea x_1 y x_2 pertenecientes a $V(\lambda)$, $f(x_1) = \lambda x_1$ y $f(x_2) = \lambda x_2$ implican

$$f(x_1 + x_2) = f(x_1) + f(x_2) = \lambda x_1 + \lambda x_2 = \lambda(x_1 + x_2)$$

y para todo α de K

$$f(\alpha x_1) = \alpha f(x_1) = \alpha(\lambda x_1) = \lambda(\alpha x_1)$$

luego

$$[x_1 \in V(\lambda),\ x_2 \in V(\lambda)] \Rightarrow [x_1 + x_2 \in V(\lambda),\ \alpha x_1 \in V(\lambda)]$$

en consecuencia, $V(\lambda)$ es un subespacio vectorial de E.

Resulta de lo anterior que si $x \in V(\lambda)$, $f(x) = \lambda x \in V(\lambda)$, luego $V(\lambda)$ es estable por f.

TEOREMA 1.—*Dado un endomorfismo f de un espacio vectorial* E *sobre el cuerpo conmutativo* K:

1. *A todo vector propio $x \neq 0$ de f corresponde un valor propio único λ, llamado valor propio asociado a x.*
2. *A todo valor propio λ de f corresponde un subespacio vectorial* $V(\lambda)$ *de* E, *que está descrito por los vectores x de* E *que verifican $f(x) = \lambda x$, se le llama* el subespacio propio asociado a λ; $V(\lambda)$ *es distinto de* $\{0\}$ *y es estable por f.*

OBSERVACIONES

1. En la definición de un vector propio x no hemos introducido la condición $x \neq 0$; esto nos ha podido sorprender, puesto que $x = 0$ verifica siempre $f(0) = \lambda 0$ cualquiera que sea λ; si se hubiera obrado así $V(\lambda)$ no contendría el cero y no sería un subespacio vectorial de E.
2. Si $\lambda = 0$ es un valor propio, existe $x \neq 0$ tal que $f(x) = 0$, luego $V(0) = \operatorname{Ker} f \neq \{0\}$.
3. Si se designa por $V(\mu)$ el conjunto de los vectores x verificando $f(\mu x) = \mu x$, la condición $V(\mu) = \{0\}$ significa que μ no es un valor propio de f.

b) Propiedades de los valores propios y de los vectores propios de un endomorfismo

Si λ es valor propio de f existe $x \neq 0$ de E tal que

$$f(x) = \lambda x \Leftrightarrow (f - \lambda e)(x) = 0$$

entonces $\operatorname{Ker}(f - \lambda e) \neq \{0\}$, es decir, $f - \lambda e$, no es inyectiva; en consecuencia, tampoco inversible. Si se supone E de *dimensión finita,* si $f - \lambda e$ no es inversible, tampoco es inyectiva (ver § 141, corolario 2 del teorema 7), de donde:

TEOREMA 2.—*Si f es un endomorfismo de* E, *espacio vectorial* de dimensión finita *sobre* K, *para* $\lambda \in K$, *las propiedades siguientes son equivalentes:*
1. *λ es un valor propio de f.*
2. *$f - \lambda e$ no es inversible.*

Consideremos dos valores propios distintos $\lambda_1 \neq \lambda_2$, busquemos x perteneciendo a $V(\lambda_1) \cap V(\lambda_2)$

$$f(x) = \lambda_1 x = \lambda_2 x \Rightarrow (\lambda_1 - \lambda_2)x = 0 \Rightarrow x = 0.$$

TEOREMA 3.—*Los subespacios vectoriales* $V(\lambda_1)$ *y* $V(\lambda_2)$ *asociados a dos valores propios distintos de un endomorfismo f, sólo tienen de común el vector nulo.*

Sea $\lambda_1, \lambda_2, \ldots, \lambda_m$, m valores propios *distintos dos a dos*; asociemos a λ_i un x_i no nulo de $V(\lambda_i)$ $(1 \leq i \leq m)$, entonces decimos que $x_1, \ldots, x_m$ son independientes.

El resultado es verdadero para $m = 1$, supongámosle verdadero para $m - 1$. De la relación

$$\alpha_1 x_1 + \ldots + \alpha_m x_m = 0$$

se deduce

$$f(\alpha_1 x_1 + \ldots + \alpha_m x_m) = \alpha_1 \lambda_1 x_1 + \ldots + \alpha_m \lambda_m x_m = 0$$

pero se tiene también, multiplicando la relación $\alpha_1 x_1 + \ldots + \alpha_m x_m = 0$ por λ_1,

$$\lambda_1 \alpha_1 x_1 + \ldots + \lambda_1 \alpha_m x_m = 0$$

de donde por sustracción

$$\alpha_2(\lambda_2 - \lambda_1)x_2 + \ldots + \alpha_m(\lambda_m - \lambda_1)x_m = 0$$

siendo las λ_i distancias dos a dos, la hipótesis de inducción da

$$(i = 2, \ldots, m) \qquad \alpha_i(\lambda_i - \lambda_1) = 0 \Rightarrow \alpha_i = 0$$

sustituyendo en la primera relación se obtiene $\alpha_1 = 0$, de donde:

TEOREMA 4. — *Siendo f un endomorfismo de* E, *espacio vectorial sobre* K *que admite m valores propios distintos dos a dos, $\lambda_1, \ldots, \lambda_m$, la familia (x_i) $(1 \leq i \leq m)$, donde x_i es un vector propio no nulo asociado a λ_i, es libre.*

COROLARIO 1. — *Si* dim E $= n$ *todo endomorfismo de* E *tiene al menos n valores propios distintos dos a dos.*

COROLARIO 2. — *Si $\lambda_1, \ldots, \lambda_m$ son dos a dos distintos el subespacio* $V(\lambda_1) + \ldots + V(\lambda_m)$ *es suma directa de* $V(\lambda_1), \ldots, V(\lambda_m)$.

En efecto, si $x_i \in V(\lambda_i)$ $(1 \leq i \leq m)$, y si

$$x_1 + x_2 + \ldots + x_m = 0$$

cada x_i es nulo, sino $x_{i_1}, \ldots, x_{i_l}$ $(l \leq m)$ no nulos formarían una familia ligada, lo que es imposible según el teorema 4.

Se deduce por sustracción que para todo x de $V(\lambda_1) + \ldots + V(\lambda_m)$

$$x = x_1 + \ldots + x_i + \ldots + x_m = x'_1 + \ldots + x'_i + \ldots + x'_m$$

x_i y x'_i pertenecientes a $V(\lambda_i)$, implica $x_i = x'_i$ para todo i, luego (§ 132, *b*)

$$V(\lambda_1) + \ldots + V(\lambda_m) = V(\lambda_1) \oplus \ldots \oplus V(\lambda_m).$$

EJERCICIO

Si λ es valor propio de f, demostrar que λ^k es valor propio de f^k cualquiera que sea el entero natural k; si, además, f es inversible λ^k es valor propio de f^k cualquiera que sea el entero racional k.

218. Polinomio característico de un endomorfismo de E, de dimensión n sobre K

En el párrafo precedente hemos definido y estudiado valores propios y vectores propios de un endomorfismo; pero no sabemos si existe. Vamos a estudiar este problema de la existencia de valores propios de f limitándonos al caso en que E es de dimensión finita sobre K, por mediación de $A = M(f)$.

a) Valores propios y vectores propios de una matriz cuadrada

Escogida una base en E de dimensión n sobre K, la biyección $f \to A = M(f)$ nos permite extender las nociones definidas para f a toda matriz de $M_n(K)$: *los valores propios, los vectores propios de* $A = M(f)$ *son por definición los valores propios y los vectores propios de* f.

Si λ y x son un valor propio y un vector propio asociados de f —por lo tanto, de $A = M(f, (a_i))$— si se pone $X = (x, (a_i))$ se tendrá

$$AX = \lambda X$$

el teorema 2 puede ser completado así (pues $M(e) = I_n$):

TEOREMA 2'. — *Siendo* A *una matriz de* $M_n(K)$, *para todo* λ *de* K *las propiedades siguientes son equivalentes:*

1. λ *es valor propio de* A.
2. $A - \lambda I_n$ *no es inversible.*
3. $\det (A - \lambda I_n) = 0$.

OBSERVACION

A es una matriz cuadrada de orden n, por definición todas las matrices $B = P^{-1}AP$ (P matriz cuadrada de orden n, inversible) tienen los mismos valores propios y los mismos vectores propios que A: éstos son los de f asociada a A en una base particular.

b) Polinomio característico de una matriz o de un endomorfismo

Sea f un endomorfismo de E de dimensión n sobre K y $A = M(f, (a_i)) = (\alpha_i^j)$. Según la tercera condición del teorema 2', λ será un valor propio si y sólo si

$$\det (A - \lambda I_n) = \begin{vmatrix} \alpha_1^1 - \lambda & \alpha_2^1 \ldots\ldots\ldots & \alpha_n^1 \\ \alpha_1^2 & \alpha_2^2 - \lambda \ldots & \alpha_n^2 \\ \vdots & \vdots & \vdots \\ \alpha_1^n & \alpha_2^n \ldots\ldots\ldots & \alpha_n^n - \lambda \end{vmatrix} = 0.$$

Sea X una indeterminada, α_i^j, X son los elementos de K[X], luego del cuerpo conmutativo K(X) (ver § 171); podemos, pues, considerar el determinante

$$p_A(X) = \det (A - XI_n) = \begin{vmatrix} \alpha_1^1 - X & \alpha_2^1 \ldots\ldots\ldots & \alpha_n^1 \\ \alpha_1^2 & \alpha_2^2 - X \ldots & \alpha_n^2 \\ \vdots & \vdots & \vdots \\ \alpha_1^n & \alpha_n^2 \ldots\ldots\ldots & \alpha_n^n - X \end{vmatrix}.$$

Este determinante det $(A - XI_n)$ es un elemento del cuerpo $K(X)$; de hecho, es un elemento de $K[X]$, ordenémosle en X; cada término contiene un elemento y uno sólo de cada columna; es, pues, de grado a lo sumo igual a n en X. Podemos, además, observar que todo término distinto del término diagonal, es decir, producto de los elementos diagonales de la matriz $A - XI_n$ es a lo sumo de grado $n - 2$. En efecto, si el término considerado como factor α_i^j $(i \neq j)$, no contiene como factor ni $\alpha_i^i - X$, ni $\alpha_j^j - X$, teniendo cada término como factor *un elemento y uno sólo* de cada línea y de cada columna. Por lo tanto, los términos del polinomio $p_A(X)$ de grado n o $n - 1$ provienen del término principal

$$(\alpha_1^1 - X)(\alpha_2^2 - X)\ldots(\alpha_n^n - X) = (-1)^n X^n + (-1)^{n-1}(\alpha_1^1 + \ldots + \alpha_n^n)X^{n-1} \ldots + \alpha_1^1\alpha_2^2, \ldots, \alpha_n^n.$$

Por otra parte, el término constante del polinomio $p(X)$ si es $p(0)$ se tiene

$$p_A(X) = \det (A - XI_n) = (-1)^n X^n + (-1)^{n-1}(\alpha_1^1 + \ldots + \alpha_n^n)X^{n-1} + \ldots + \det A$$

$p_A(X)$ es, pues, de grado n, se llama polinomio característico de la matriz A.

Sea B una matriz semejante a A, es decir, tal que $B = P^{-1}AP$, siendo P una matriz inversible de orden n con elementos en K, tenemos en el anillo de las matrices con coeficientes en el cuerpo $K(X)$

$$B - XI_n = P^{-1}AP - X(P^{-1}I_nP) = P^{-1}AP - P^{-1}(XI_n)P = P^{-1}(A - XI_n)P$$

pues

$$I_n = P^{-1}I_nP,$$

de donde

$$p_B(X) = \det (B - XI_n) = (\det P^{-1}) \det (A - XI_n)(\det P) = \det (A - XI_n) = p_A(X).$$

Sea f un endomorfismo de E de dimensión n sobre K, si $A = M(f,(a_i))$ toda matriz asociada a f en una base cualquiera es $B = P^{-1}AP$, de donde:

TEOREMA Y DEFINICIÓN 5. — *Si f es un endomorfismo del espacio vectorial* E *de dimensión n sobre* K, A *una matriz asociada a f, el polinomio*

$$p_A(X) = \det (A - XI_n)$$

es invariante cuando se reemplaza A *por una matriz semejante, se dice que $p_A(X)$ es el* polinomio característico de $A \in M_n(K)$ *o el polinomio característico de $f \in \mathfrak{L}(E)$, se le designa también $p_f(X)$.*

De lo cual resulta que si se escribe

$$A = (\alpha_i^j) = M(f), \; p_f(X) = p_A(X) = a_nX^n + a_{n-1}X^{n-1} + \ldots + a_0.$$

Los coeficientes $a_n, \ldots, a_0$ elementos de K se expresan por medio de polinomios en α_i^j hemos visto que

$$a_n = (-1)^n, \qquad a_{n-1} = (-1)^{n-1}(\alpha_1^1 + \ldots + \alpha_n^n), \qquad a_0 = \det A.$$

Según lo que acabamos de ver los coeficientes son invariantes cuando se reemplaza A por una matriz semejante cualquiera, ***sólo dependen del endomorfismo*** f; $(-1)^{n-1}a_{n-1}$ se llama la *traza* de f, o de A, y se escribe tr (f) o tr (A); por lo tanto,

$$A = M(f) = (\alpha_i^j) \Rightarrow \text{tr}\ (f) = \text{tr}\ (A) = \alpha_1^1 + \dots + \alpha_n^n$$

se tiene, pues,

$$p_f(X) = (-1)^n X^n + (-1)^{n-1}\ \text{tr}\ (f) X^{n-1} + \dots + \det\ (f).$$

c) Investigación de los valores y vectores propios

Si f es un endomorfismo de E de dimensión n sobre K, los resultados precedentes nos permiten enunciar:

TEOREMA 6. — *Siendo f un endomorfismo de un espacio vectorial* E *de dimensión n sobre el cuerpo conmutativo* K, *los valores propios de f son las raíces de su polinomio característico. Existen n a lo sumo.*

Si K *es algebraicamente cerrado, f posee n valores propios distintos o confundidos.*

Habiendo escogido una base (a_i) de E si se pone

$$A = M(f,\ (a_i)) = (\alpha_i^j).$$

Debemos resolver la ecuación algebraica de grado n, $p_A(\lambda) = 0$.

Esta última puede que no tenga ninguna raíz en K. Tomemos, por ejemplo, $K = \mathbf{R}$ y

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

$$p_A(\lambda) = \begin{vmatrix} a-\lambda & b \\ c & d-\lambda \end{vmatrix} = \lambda^2 - (a+d)\lambda + ad - bc = 0.$$

Si $(a+d)^2 - 4(ad-bc) = (a-d)^2 + 4bc < 0$, la matriz A y el endomorfismo asociado A (endomorfismo de E de dimensión 2 sobre **R**) no tiene valor propio en **R**; luego no hay vectores propios no nulos.

Hemos dicho que todo polinomio $p \in K[X]$, de grado n, tenía n raíces en su cuerpo de las raíces Δ (ver § 193, *a*) y ej. 348, capítulo 11); nos tomaremos la libertad de hablar de las n raíces de p_f especificando si están o no en K; por ejemplo, en el ejemplo precedente las dos raíces de p_A están en **C** y no en **R**.

Supongamos que existe $\lambda \in K$, $p_A(\lambda) = 0$, un vector propio $x = \sum_{i=1}^{n} x^i a_i$ asociado a λ estará determinado por la ecuación

$$f(x) = \lambda x.$$

Es decir, en la base (a_i) sus coordenadas serán soluciones del sistema

$$\begin{cases} (\alpha_1^1 - \lambda)x^1 + \alpha_2^1 x^2 + \dots + \alpha_n^1 x^n = 0 \\ \alpha_1^2 x^1 + (\alpha_2^2 - \lambda)x^2 + \dots + \alpha_n^2 x^n = 0 \\ \vdots \qquad \vdots \qquad \vdots \\ \alpha_1^n x^1 + \alpha_2^n x^2 + \dots + (\alpha_n^n - \lambda)x^n = 0 \end{cases}$$

Ahora bien, el determinante de este sistema homogéneo de n ecuaciones con n incógnitas es $p_A(\lambda) = 0$, luego existen otras soluciones, además de la solución trivial (ver § 180). Este sistema es de rango $r < n$, encontraremos soluciones $x = (x^1, \dots, x^n)$ que describen un subespacio vectorial de E, que no es otro si no el $V(\lambda)$ y que es de dimensión $n - r$. En particular r puede ser nulo, $V(\lambda)$ entonces es de dimensión n. Como siempre hay $x \neq 0$ en $V(\lambda)$ vemos que $1 \leq \dim V(\lambda) \leq n$. Si λ es una raíz en K de orden k de p_f podemos precisar más la dimensión de $V(\lambda)$, supongamos $\dim V(\lambda) = h$, sea $(a_1, \dots, a_h)$ una base de $V(\lambda)$, completémosla por $(a_{h+1}, \dots, a_n)$ para obtener una base de E; a_i $(1 \leq i \leq h)$ es un vector propio no nulo asociado a λ tendremos, pues, con los vectores columnas de A las imágenes por f de los vectores de la base

$$A = M(f, (a_i)) = \left(\begin{array}{c|c} \lambda I_h & A' \\ \hline 0 & A'' \end{array}\right)$$

el bloque A'' que es una matriz cuadrada de orden $n - h$. Obtenemos (ver § 170, *b*)

$$p_A(X) = (\lambda - X)^h \det (A'' - XI_{n-h}) = (\lambda - X)^h q(X)$$

q que es un polinomio de de K[X], luego es imposible que $h > k$, puesto que λ es de orden k, luego

TEOREMA 7. — *Siendo f un endomorfismo de* E *de dimensión n sobre* K *y si su polinomio característico p_f admite en* K *una raíz múltiple λ de orden k, se tiene*

$$1 \leq \dim V(\lambda) \leq k.$$

219. Diagonalización

a) DEFINICIÓN. — *Se dice que un endomorfismo f del espacio vectorial* E *de dimensión n sobre* K *es* diagonalizable *si existe una base de* E *tal que la matriz asociada a f relativamente a esta base sea diagonal.*

Se dice que A *de* $M_n(K)$ *es diagonalizable si existe una matriz cuadrada inversible* P *de orden n tal que* $P^{-1}AP$ *sea diagonal.*

Se ve que las condiciones:

1. f de $\mathcal{L}_k(E)$ es diagonalizable.
2. $A = M(f, (a_i))$ es diagonalizable

son las mismas: P es la matriz de cambio de la base (a_i) a la base (b_i) de E tal que $M(f, (b_i))$ sea diagonal.

b) **Condición necesaria y suficiente de diagonalización**

Sea f un endomorfismo de E de dimensión n sobre K tal que en la base (a_i) la matriz A asociada a f sea diagonal.

$$A = \begin{pmatrix} \mu_1 & & & & O \\ & \ddots & & & \\ & & \mu_i & & \\ & & & \ddots & \\ O & & & & \mu_n \end{pmatrix}$$

tenemos

$$p_A(X) = (\mu_1 - X) \dots (\mu_i - X) \dots (\mu_n - X)$$

luego $\mu_1, \dots, \mu_n$ son los valores propios de A (o de f). Por lo tanto, las raíces de $p_A(X)$ están todas en K. Por otra parte, para todo i de $[1, n]$

$$f(a_i) = \mu_i a_i$$

luego tenemos que la base de E está formada de vectores propios no nulos..

Recíprocamente supongamos que se pueda hallar una base (a_i) de vectores propios, es evidente que respecto a esta base la matriz asociada a f es diagonal de donde:

Teorema 8. — *Un endomorfismo f de* E, *espacio vectorial de dimensión n sobre* K, *es diagonalizable si y sólo si es posible hallar una base de* E *formada de vectores propios.*

Acabamos de ver que si f es diagonalizable, sus n valores propios (distintos o confundidos) están en K, vamos a ver que se puede dar al teorema 8 una forma más precisa haciendo intervenir el orden de multiplicidad de las raíces de p_f; supongamos que λ_i sea de orden k_i, tendremos en K[X]

$$p_f(X) = (-1)^n(X - \lambda_1)^{k_1} \dots (X - \lambda_m)^{k_m}$$
$$k_1 + \dots + k_m = n, \qquad (\lambda_i \in K,\ i = 1, \dots, m).$$

Supongamos f diagonizable, una base donde $M(f)$ es diagonal está formada de vectores propios, ordenemos los elementos de esta base de manera de obtener primero los vectores propios relativos a λ_1, a continuación los relativos a λ_2 y, en último lugar, los relativos a λ_m.

En esta base

$$A = M(f) = \begin{pmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & 0 & \\ & & \lambda_1 & & & & & \\ & & & \lambda_2 & & & & \\ & & & & \ddots & & & \\ & & & & & \lambda_2 & & \\ & & & & & & \ddots & \\ & 0 & & & & & \lambda_m & \\ & & & & & & & \ddots \\ & & & & & & & & \lambda_m \end{pmatrix};$$

por lo tanto, $E = V(\lambda_1) \oplus \ldots \oplus V(\lambda_m)$; de donde

$$\dim E = \dim V(\lambda_1) + \ldots + \dim V(\lambda_m) = n$$

según el teorema 7: $\dim V(\lambda_i) \leq k_i$ $(i = 1, \ldots, m)$; la relación $k_1 + \ldots + k_m = n$ nos da

(1) $\qquad (i = 1, 2, \ldots, m) \qquad \dim V(\lambda_i) = k_i.$

Recíprocamente las n raíces de p_f estando en K, supongamos que la condición (1) anterior esté satisfecha, la relación $k_1 + \ldots + k_m = n$ implica

$$\dim [V(\lambda_1) \oplus \ldots \oplus V(\lambda_m)] = n$$

luego $V(\lambda_1) \oplus \ldots \oplus V(\lambda_m)$ incluido en E, y teniendo la misma dimensión E, es igual a E; la reunión de las bases de $(V(\lambda_i))$ $(1 \leq i \leq m)$ es una base de E formada de vectores propios y f es diagonizable, de donde:

TEOREMA 8'. — *Un endomorfismo f de* E, *espacio vectorial de dimensión n sobre* K, *es diagonalizable si y sólo si:*

1. *El polinomio característico p_f tiene sus n raíces (distintas o confundidas) en* K.
2. *Para cada raíz λ_i de p_f, de orden k_i,*

$$\dim V(\lambda_i) = k_i.$$

Todas las raíces de p_j están en K; se ve entonces que la segunda condición equivale a decir: cada $V(\lambda_i)$ tiene la mayor dimensión posible. El isomorfismo $f \to M(f)$ nos permite enunciar:

COROLARIO. — *Una matriz* A *de* $M_n(K)$ *es semejante a una matriz diagonal si y sólo si:*

1. *El polinomio característico p_A tiene sus n raíces en* K.
2. *Para cada raíz λ_i de p_A, de orden k_i,*

$$\dim V(\lambda_i) = k_i.$$

c) Condición suficiente de diagonalización

Supongamos que todas las raíces de $p_f(X)$ son elementos de K, $\lambda_1, \ldots, \lambda_n$ y que sean *dos a dos distintas*, existe una familia de vectores (x_i)

$$(i = 1, 2, \ldots, n) \qquad f(x_i) = \lambda_i x_i, \qquad x_i \neq 0$$

luego (x) $(1 \leq i \leq n)$ es una familia libre de vectores propios (ver § 217, t. 4), tiene n elementos, es, pues, una base de E formada de vectores propios, de donde:

TEOREMA 9. — *Si f, endomorfismo de* E, *de dimensión n sobre* K, *o* A *elemento de* $M_n(K)$ *posee n valores propios todos distintos en* K, *f y* A *son diagonalizables.*

En particular si K es *algebraicamente cerrado,* por ejemplo, si K = C, todas las raíces del polinomio característico están en K, son dos a dos distintas y si son simples, de donde:

COROLARIO. — *Siendo f un endomorfismo de* E, *de dimensión n sobre* K *y* A *un elemento de* $M_n(K)$, *si* K *es algebraicamente cerrado, para que f y* A *sean diagonalizables* es suficiente *que todas las raíces de su polinomio característico sean simples.*

OBSERVACION

El teorema 9 o el corolario dan sólo una condición suficiente de diagonalización: todos los valores propios de f estando en K, f puede ser muy bien diagonalizable incluso si p_f tiene raíces múltiples. Es suficiente considerar el ejemplo trivial $A = I_n$, cuyo polinomio característico es $(1 - X)^n$ (ver igualmente ejemplo 3 más abajo).

EJEMPLOS Y EJERCICIOS

1. Determinar los valores propios y los vectores propios de la matriz

$$A = \begin{pmatrix} 2 & 0 & 4 \\ 3 & -4 & 12 \\ 1 & -2 & 5 \end{pmatrix} \in M_3(\mathbf{R}) \qquad \text{(M.G.P.)}$$

$$p_A(X) = \begin{vmatrix} 2-X & 0 & 4 \\ 3 & -4-X & 12 \\ 1 & -2 & 5-X \end{vmatrix} = -X(X-1)(X-2).$$

Los tres valores propios pertenecen a **R**, los tres son simples, la matriz A es diagonalizable. Busquemos los vectores propios:

a) $\lambda_1 = 0$ las coordenadas (x_1, x_2, x_3) de un vector propio en relación a la base (a_i) tal que $A = M(f, (a_i))$ verifican

$$\begin{cases} 2x_1 + 4x_3 = 0 \\ 3x_1 - 4x_2 + 12x_3 = 0 \\ x_1 - 2x_2 + 5x_3 = 0 \end{cases} \Rightarrow \frac{x_1}{-4} = \frac{x_2}{3} = \frac{x_3}{2}.$$

b) $\lambda_2 = 1$ tendremos

$$\begin{cases} x_1 + 4x_3 = 0 \\ 3x_1 - 5x_2 + 12x_3 = 0 \\ x_1 - 2x_2 + 4x_3 = 0 \end{cases} \Rightarrow \frac{x_1}{-4} = \frac{x_2}{0} = \frac{x_3}{1}$$

c) $\lambda_3 = 2$ tendremos

$$\begin{cases} x_1 + 4x_3 = 0 \\ 3x_1 - 6x_2 + 12x_3 = 0 \\ x_1 - 2x_2 + 3x_3 = 0 \end{cases} \Rightarrow \frac{x_1}{2} = \frac{x_2}{1} = \frac{x_3}{0}.$$

Los vectores propios relativos a los vectores propios 0, 1, 2 son, pues, respectivamente (α_1, α_2, α_3 reales no nulos cualesquiera),

$$\begin{aligned} \lambda_1 &= 0 : \alpha_1(-4a_1 + 3a_2 + 2a_3) \\ \lambda_2 &= 1 : \alpha_2(-4a_1 + a_3) \\ \lambda_3 &= 2 : \alpha_3(2a_1 + a_2) \end{aligned}$$

Cada uno de ellos describe un subespacio vectorial de 1 dimensión; estos tres vectores son independientes y forman una base de E. Designemos por (b_i) la base obtenida tomando, por ejemplo, $\alpha_1 = \alpha_2 = \alpha_3 = 1$, la matriz de paso de la base (a_i) a la base (b_i) será (ver § 160)

$$\mathrm{P} = \begin{pmatrix} -4 & -4 & 2 \\ 3 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}.$$

Sabemos que $\mathrm{B} = \mathrm{M}(f, (b_i))$ es diagonal y tiene por elementos diagonales en este orden 0, 1, 2. Se podrá a manera de ejercicio calcular P^{-1} y verificar que

$$\mathrm{B} = \mathrm{P}^{-1}\mathrm{AP} = \frac{1}{2}\begin{pmatrix} -1 & 2 & -4 \\ 2 & -4 & -10 \\ 3 & -4 & 12 \end{pmatrix}\begin{pmatrix} 2 & 0 & 4 \\ 3 & -4 & 12 \\ 1 & -2 & 5 \end{pmatrix}\begin{pmatrix} -4 & -4 & 2 \\ 3 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

2. Determinar los valores propios y los vectores propios de

$$\mathrm{A} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}.$$

a) Considerando $\mathrm{A} \in \mathrm{M}_3(\mathbf{R})$, *b*) considerando $\mathrm{A} \in \mathrm{M}_3 \in (\mathbf{C})$; se obtiene $\lambda_1 = 0$, $\lambda_2 = j - 1$, $\lambda_3 = j^2 - 1$ $(j^3 = 1)$, luego A es diagonalizable en $\mathrm{M}_3(\mathbf{C})$ y no en $\mathrm{M}_3(\mathbf{R})$.

3. Determinar los valores propios y los vectores propios de

$$\mathrm{A} = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \in \mathrm{M}_3(\mathbf{R}).$$

Se obtiene $p_\mathrm{A}(\mathrm{X}) = -(\mathrm{X} - 1)(\mathrm{X} + 2)^2$; buscando los vectores propios se halla que $\dim \mathrm{V}(-2) = 2$; por tanto, A es diagonalizable. Veremos en el capítulo 15 que este resultado no es casual: *toda matriz cuadrada simétrica real tiene todos sus valores propios reales y es siempre diagonalizable.*

Por otra parte, este ejemplo muestra que todos los valores propios estando en K, el hecho de que estas raíces sean simples es una *condición suficiente, pero no necesaria,* para que A sea diagonalizable.

4. Determinar los valores propios y los vectores propios de

$$A = \begin{pmatrix} -4 & 0 & -2 \\ 0 & 1 & 0 \\ 5 & 1 & 3 \end{pmatrix} \in \mathbf{M}_3(\mathbf{R}).$$

Se obtiene $p_A(X) = -10(X-1)^2(X+2)$ y dim $V(1) = \dim V(-2) = 1$, luego A no es diagonalizable.

220. Reducción a la forma triangular

Sea f un endomorfismo de E de dimensión n sobre K, $A = M(f, (a_i))$; si f y A no son diagonalizables, pero si f tiene todos sus valores propios en K (y en particular si K es *algebraicamente cerrado*), podemos hallar una base (b_i) de E tal que

$$B = P^{-1}AP = M(f, (b_i))$$

sea triangular, es decir, vamos a demostrar el resultados siguiente:

a) TEOREMA 10. — *Para todo endomorfismo f de* E *de dimensión n sobre* K, *tal que el polinomio característico de f tenga todas sus raíces en* K, *existe una base de* E *para la cual la matriz* B *asociada a f en esta base sea triangular, los elementos diagonales de* B *son los valores propios de f.*

Observemos primero que si en la base $(e_1, \ldots, e_n)$, $M(f) = (\alpha_i^j)$ en la base $(e_n, \ldots, e_1)$ se tendrá: $M(f) = (\alpha_{n-i}^{n-j})$; en efecto, la nueva base (e'_i) es tal que $e'_i = e_{n-i}$ para todo i de $[1, n]$, luego

$$f(e'_i) = f(e_{n-i}) = \sum_{k=1}^{n} \alpha_{n-i}^k e_k = \sum_{k=1}^{n} \alpha_{n-i}^k e'_{n-k}$$

$$= \sum_{j=1}^{n} \alpha_{n-i}^{n-j} e'_j = \sum_{j=1}^{n} \alpha'^j_i e'_j .$$

Ahora bien $(\alpha'^j_i) = (\alpha_{n-i}^{n-j})$ se deduce de (α_i^j) por una "*simetría respecto al centro*" de la matriz (α_i^j). *Luego si* $M(f)$ *es triangular superior relativamente a la base* $(e_1, \ldots, e_n)$, $M(f)$ *es triangular inferior relativamente a la base* $(e_n, \ldots, e_1)$: No es, pues, restrictivo el buscar una matriz B triangular superior.

El teorema es trivial para $n = 1$; supongámoslo demostrado para $n - 1$; el polinomio característico de f teniendo todas sus raíces en K, sea λ_1 una de ellas y b_1 un vector propio no nulo asociado a λ_1

$$f(b_1) = \lambda_1 b_1, \qquad b_1 \neq 0.$$

Designemos por Kb_1 el subespacio vectorial de dimensión uno, engendrado por b_1, y sea (b'_i) $(1 \leq i \leq n)$ una base cualquiera de un suplemento cualquiera E' de Kb_1, $(b_1, b'_2, \ldots, b'_n)$ es una base de E y relativamente a esta base la matriz de f es

$$(1) \qquad B' = \begin{pmatrix} \lambda_1 & \beta_2'^1 & \ldots & \beta_n'^1 \\ 0 & \beta_2'^2 & \ldots & \beta_n'^2 \\ \vdots & \vdots & & \\ \vdots & \vdots & & \\ 0 & \beta_2'^n & \ldots & \beta_n'^n \end{pmatrix} = \left(\begin{array}{c|c} \lambda_1 & \beta_2'^1 \ldots \beta_n'^1 \\ \hline 0 & \\ \vdots & C' \\ 0 & \end{array}\right).$$

Intentemos interpretar el bloque C' que es un elemento de $M_{n-1}(K)$, tenemos

$$(i = 2, \ldots, n) \qquad f(b'_i) = \beta_i'^1 b_1 + \beta_i'^2 b'_2 + \ldots + \beta_i'^n b'_n$$

en general los $n - 1$ escalares $\beta_2'^1, \ldots, \beta_n'^1$ no son todos nulos, E' *no es estable por f*. Pero consideremos la proyección π de E sobre E' paralelamente a Kb_1 y el endomorfismo $g = \pi \circ f$ de E, tendremos

$$(i = 2, \ldots, n) \qquad g(b'_i) = \beta_i'^2 b'_2 + \ldots + \beta_i'^n b'_n$$

luego E' *es estable por g*, designemos por g' la aplicación inducida por g sobre E', tendremos $(2 \leq i \leq n)$

$$M(g', (b'_i)) = C'$$

para aplicar la hipótesis de recurrencia nos hace falta demostrar que g' (o C') tiene todos sus valores propios en K; ahora bien, la fórmula (1) muestra desarrollando det $(B' - XI_n)$ respecto a su primera columna que

$$p_f(X) = p_{B'}(X) = (\lambda_1 - X) \det (C' - XI_{n-1}) = (\lambda_1 - X) p_{g'}(X)$$

luego el conjunto de los valores propios de g' (o C') *es el conjunto de los valores propios de f menos una vez el de* λ_1: todos los valores propios de g' están en K la hipótesis de recurrencia puede serle aplicada, existe, pues, una base b_i $(2 \leq i \leq n)$ de E' tal que $C = M(g', (b_i))$ sea triangular superior, es decir,

$$C = M(g', (b_i)) = \begin{pmatrix} \beta_2^2 & \beta_3^2 & \ldots\ldots & \beta_n^2 \\ 0 & \beta_3^3 & \ldots\ldots & \beta_n^3 \\ 0 & 0 & \ldots\ldots & \beta_n^4 \\ \vdots & \vdots & & \\ 0 & 0 & \ldots\ldots & \beta_n^n \end{pmatrix}.$$

Si la relación (1) es válida cualquiera que sea la base escogida en E', relativamente a la base (b_i) $(1 \leq i \leq n)$ de E, tendremos designado por $\beta_j^1 (2 \leq j \leq n)$ la coordenada de $f(b_j)$ sobre b_1

$$B = M(f, (b_i)) = \left(\begin{array}{c|c} \lambda_1 & \beta_2^1 \ldots \beta_n^1 \\ \hline 0 & \\ \vdots & C \\ 0 & \end{array}\right) = \begin{pmatrix} \lambda_1 & \beta_2^1 & \ldots\ldots & \beta_n^1 \\ 0 & \beta_2^2 & \ldots\ldots & \beta_n^2 \\ 0 & 0 & \ldots\ldots & \beta_n^3 \\ \vdots & \vdots & & \\ 0 & 0 & \ldots\ldots & \beta_n^n \end{pmatrix}$$

luego B es triangular superior. Se ve inmediatamente que (ver § 170, *b*)

$$p_f(X) = p_B(X) = (\lambda_1 - X)(\beta_2^2 - X) \dots (\beta_n^n - X)$$

los valores propios de f son, pues, los elementos diagonales de B.

COROLARIO. — *Siendo* A *un elemento de* $M_n(K)$ *cuyo polinomio característico tiene todas sus raíces en* K, *existe una matriz* B *triangular semejante a* A.

El teorema 10 y su corolario son evidentemente aplicables si K es *algebraicamente cerrado, y, en particular, si* $K = \mathbf{C}$.

b) **Teorema de Cayley-Hamilton**

Siendo q un polinomio de K[X] representado por

$$q(X) = a_0 + a_1X + \dots + a_kX^k$$

se puede definir $q(x)$ para todo elemento x de un superanillo L de K no forzosamente conmutativo *a condición de que* x *permute con todo elemento de* K. Es el caso de f, elemento de $\mathfrak{L}_k(E)$, E espacio vectorial de dimensión n sobre K, o de A elemento de $M_n(K)$ (ver § 185, *b*, observación y ejemplo 5), se tiene

$$q(f) = a_0e + a_1f + \dots + a_kf^k \in \mathfrak{L}_K(E)$$

con

$$e = \mathrm{id}_E = f^0, \qquad f^1 = f, \qquad f^h = f^{h-1} \circ f \qquad (\text{para} \quad h > 1)$$

y

$$q(A) = a_0I_n + a_1A + \dots + a_kA^k \in M_n(K)$$

$q(b)$ y $q(A)$ son llamados, respectivamente, *polinomio de endomorfismo* y *polinomio de matriz*. El hecho de que f permute con todo elemento de K y la relación $f^{l+m} = f^l \circ f^m$ $(l, m \in \mathbf{N})$ implican ($q, r, s,$ elementos de K[X])

$$q = rs \Rightarrow q(f) = r(f) \circ s(f), \qquad q(A) = r(A)s(A)$$

si K es algebraicamente cerrado, tenemos en K[X]

$$p_f(X) = (-1)^n(X - \lambda_1) \dots (X - \lambda_n) \qquad (\lambda_i \in K, \; i = 1, 2, \dots, n)$$

por lo que

$$p_f(f) = (-1)^n(f - \lambda_1e) \circ \dots \circ (f - \lambda_ne) = (-1)^n(g_1 \circ \dots \circ g_n) = (-1)^ng$$

poniendo $g_i = f - \lambda_ie (i = 1, 2, \dots, n)$. Luego $g = g_1 \circ \dots \circ g_n$ es un endomorfismo de E de dimensión n sobre K.

Diremos que un endomorfismo φ *anula* un vector x (resp. un subespacio V de E) si $\varphi(x) = 0$ (resp. $\varphi(V) = \{0\}$). Nos proponemos demostrar que si K es algebraicamente cerrado g, luego $p_f(f)$, anula E; basta demostrar que g anula todo vector de una base de E. Existe una base (b_i) tal que $M(f, (b_i))$ sea triangular (t. 10), los elementos diagonales son los valores propios de f

$$B = M(f, (b_i)) = \begin{pmatrix} \lambda_1 & \beta_2^1 \dots \beta_n^1 \\ 0 & \lambda_2 \dots \beta_n^2 \\ \vdots & \vdots \quad\quad \vdots \\ 0 & 0 \quad \dots \lambda_n \end{pmatrix}.$$

Observemos primero que en el anillo $\mathfrak{L}(E)$

$$g_i \circ g_j = (f - \lambda_i e) \circ (f - \lambda_j e) = f^2 - (\lambda_i + \lambda_j)f + \lambda_i\lambda_j e$$
$$= (f - \lambda_j e) \circ (f - \lambda_i e) = g_j \circ g_i.$$

Demostremos que $G_i = g_1 \circ \dots \circ g_i$ anula $b_1, b_2, \dots, b_i$, el resultado es trivial para $i = 1$, pues según la forma de B

$$g_1(b_1) = (f - \lambda_1 e)(b_1) = f(b_1) - \lambda_1 b_1 = 0$$

supongamos, hipótesis de recurrencia, que este resultado sea cierto para

$$G_i = g_1 \circ g_2 \dots \circ g_i = G_{i-1} \circ g_i = g_i \circ G_{i-1}$$

puesto que G_{i-1} anula $b_1, \dots, b_{i-1}$ ocurrirá lo mismo para G_i. Por otro lado, por la forma de B

$$g_i(b_i) = (f - \lambda_i e)(b_i) = f(b_i) - \lambda_i b_i = \beta_i^1 b_1 + \dots + \beta_i^{i-1} b_{i-1}$$

en consecuencia,

$$G_i(b_i) = G_{i-1}[g_i(b_i)] = \sum_{j=1}^{i-1} \beta_i^j G_{i-1}(b_j) = 0$$

según la hipótesis de recurrencia.

El teorema es, pues, verdadero de 1 a n y $p_f(f) = (-1)^n G_n = (-1)^n g$ anula $b_1, \dots, b_n$; por lo tanto, E, es decir,

$$p_f(f) = 0.$$

TEOREMA DE CAYLEY-HAMILTON. — *Para todo endomorfismo f de un espacio vectorial de dimensión n sobre* K *y todo elemento* A *de* $M_n(K)$, *tales que sus polinomios característicos tengan todas sus raíces en* K *se tiene*

$$p_f(f) = 0, \qquad p_A(A) = 0.$$

Este teorema es cierto para todo endomorfismo de $\mathfrak{L}_k(E)$ y toda matriz de $M_f(K)$ si K es *algebraicamente cerrado*; se aplica, en particular, a $K = \mathbf{C}$, igualmente también se aplica a **R**, basta observar que

$$A \in M_n(\mathbf{R}) \subset M_n(\mathbf{C})$$

por lo tanto si $A = M(f)$, $p_A(A) = 0$ y esta igualdad implica $p_f(f) = 0$; el polinomio $p_A = p_f$ puede tener raíces complejas no reales, pertenece, sin embargo, a $\mathbf{R}[X]$.

De hecho el teorema de Cayley-Hamilton se aplica a todo cuerpo conmutativo K, basta aplicar a K y a Δ, cuerpo de las raíces de p_A (ver § 193, *a*, y ej. 348, capítulo 11), el razonamiento que acabamos de hacer sobre **R** y **C**.

Ejercicios

En los ejercicios 443 al 457 se hallará los valores propios y los vectores propios de cada matriz A *suponiendo que represente un endomorismo* f *de* $\mathbf{C}^n$ *expresado en su base canónica. Hallar eventualmente una base de vectores propios o una base en que la matriz asociada a* f *sea trianular. Calcular* P, P^{-1} *y* $P^{-1}AP$. *Como ejercicio de cálculo, constatar que cada matriz verifica su polinomio característico. Cuando* A *es diagonalizable sobre* **C** *hallar si también lo es sobre* **R** *o sobre* **Q**.

443. $\begin{pmatrix} 2 & 0 & 4 \\ 3 & -4 & 12 \\ 1 & -2 & 5 \end{pmatrix}$. **444.** $\begin{pmatrix} 3 & 1 & 0 \\ -4 & -1 & 0 \\ 4 & -8 & -2 \end{pmatrix}$. **445.** $\begin{pmatrix} 1 & 4 & 2 \\ 0 & -3 & -2 \\ 0 & 4 & 3 \end{pmatrix}$.

446. $\begin{pmatrix} -2 & -2 & 1 \\ -2 & 1 & -2 \\ 1 & -2 & -2 \end{pmatrix}$. **447.** $\begin{pmatrix} 3 & -1 & -1 \\ -6 & -1 & 2 \\ 2 & 1 & 0 \end{pmatrix}$. **448.** $\begin{pmatrix} 2 & -2 & 1 \\ 1 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix}$.

449. $\begin{pmatrix} 5 & 3 & 2 \\ 1 & 1 & 1 \\ 3 & 3 & 1 \end{pmatrix}$. **450.** $\begin{pmatrix} 3 & 1 & 1 \\ -1 & -1 & 2 \\ 9 & 3 & -2 \end{pmatrix}$. **451.** $\begin{pmatrix} 5 & -1 & 9 \\ 3 & 4 & 0 \\ 1 & 1 & 1 \end{pmatrix}$.

452. $\begin{pmatrix} 2 & 0 & 0 \\ -3 & -1 & 3 \\ 3 & 3 & -1 \end{pmatrix}$. **453.** $\begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$. **454.** $\begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix}$.

455. $\begin{pmatrix} 1 & 3 & 0 & 0 \\ 4 & 2 & 0 & 0 \\ 1 & -1 & 5 & -3 \\ 2 & 0 & 4 & -2 \end{pmatrix}$. **456.** $\begin{pmatrix} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 7 & 1 & 2 & 1 \\ -17 & -6 & -1 & 0 \end{pmatrix}$.

457. $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Generalizar a una matriz de esta forma y de orden n.

En los ejercicios siguientes, 458 al 460, se hallará los valores propios y los vectores propios en **C** *después en* **R** *y se determinará los casos en que la matriz es diagonalizable* (a, b, $c \in \mathbf{R}$, $a \neq 0$ *en el ejercicio* 458).

458. $\begin{pmatrix} 0 & a & a^2 \\ 1/a & 0 & a \\ 1/a^2 & 1/a & 0 \end{pmatrix}$. **459.** $\begin{pmatrix} 1 & a & 1 \\ 0 & 1 & b \\ 0 & 0 & c \end{pmatrix}$. **460.** $\begin{pmatrix} c & 2a & 0 \\ b & 0 & a \\ 0 & 2b & -c \end{pmatrix}$.

461. A es la matriz dada en el ejercicio 3, § 156, hallar su polinomio característico utilizando el ejercicio 224, capítulo 8. Determinar los casos en que A es diagonalizable en **C**.

462. Si a, b, c, d son números reales, hallar el polinomio característico de la matriz.

$$\begin{pmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{pmatrix}.$$

¿Es diagonalizable en **C**? Si no lo es reducirla a la forma triangular.

463.

$$A = \begin{pmatrix} 1 & 0 \\ 3/2 & 1 \end{pmatrix}.$$

a) Calcular ${}^tAA$, $A{}^tA = B$.

b) Demostrar que existe P inversible tal que $B = PCP^{-1}$, siendo C diagonal. Calcular C y P.

c) Probar que existe D tal que $D^2 = C$; calcular D.

d) Probar que existe H simétrica tal que $H^2 = B$; calcular H.

e) Probar que existe U ortogonal (es decir tal que $U^{-1} = {}^tU$) verificando $A = HU$; calcular U.

464 Si a y b son números complejos, A es la matriz de tipo $(1, n)$ cuyos elementos son todos iguales a a y B la matriz de tipo $(n, 1)$ cuyos elementos son todos iguales a b hallar los valores propios de AB.

465. Si f es un endomorfismo nilpotente de índice p de un espacio vectorial E de dimensión n sobre **K**, ¿cuáles son los valores propios de f? (V. cap. 8 ej. 192).

466. Si f es un endomorfismo de E de dimensión n sobre **C** tal que $f^2 = f$, ¿cuáles son los valores propios de f? Hallar todas las matrices de $\mathbf{M}_2(\mathbf{R})$ tales que $A^2 = A$.

467. Si f es un endomorfismo de E de dimensión n sobre **C** tal que existe un entero natural k verificando $f^k = e$, $f^{k-1} \neq e$, ¿cuáles son los valores propios de f? Caso particular: $k = 2$ (V. cap. 8 ej. 196).

468. Siendo f un endomorfismo de un espacio vectorial E de dimensión n sobre **K** se considera el endomorfismo ${}^tf$ de E*.

a) Demostrar que f y ${}^tf$ tienen el mismo polinomio característico.

b) Sea $V(\lambda)$ el subespacio de E asociado a λ, valor propio de f y $V'(\lambda)$ el subespacio de E* asociado λ, valor propio de ${}^tf$, demostrar que $V(\lambda)$ y $V'(\lambda)$ tienen la misma dimensión.

469. $A \in \mathbf{M}_n(C)$, B es su matriz adjunta, demostrar que si x es un vector propio de A asociado al valor propio λ de A, existe un valor propio μ de B tal que x sea vector propio de B asociado a μ. Se estudiará sucesivamente los casos siguientes:

a) A es inversible,

b) A es no inversible y $\lambda \neq 0$,

c) $\lambda = 0$ es raíz simple del polinomio característico de A,

d) $\lambda = 0$ es raíz múltiple de este polinomio (**École Polytechnique**)

470. Se considera la matriz A de $\mathbf{M}_n(\mathbf{C})$

$$A = \begin{pmatrix} a_0 & a_1 \dots a_{n-1} \\ a_n & a_0 \dots a_{n-2} \\ \cdot & \cdot \quad \cdot \\ \cdot & \cdot \quad \cdot \\ \cdot & \cdot \quad \cdot \\ a_1 & a_2 \dots a_0 \end{pmatrix}$$

y la matriz $P = p_{kl}$, de orden n (k índice de línea y l índice de columna) definida por

$$p_{kl} = e^{\frac{2i\pi kl}{n}}$$

a) Calcular $P\overline{P}$. Deducir que P es inversible. Determinar P^{-1}.

b) Siendo ω una raíz n-ésima de 1, demostrar que el vector de coordenadas $1, \omega, \dots, \omega^{n-1}$ es un vector propio de A. ¿Cuál es el valor propio correspondiente?

c) Calcular $P^{-1}AP$. (M.G.P., extractado)

471. Determinar los valores propios y los vectores propios de la matriz dada en el ejercicio 223 del capítulo 8 (se podrá utilizar el resultado de la pregunta **c**) de este ejercicio).

472. Se considera el endomorfismo estudiado en el ejercicio 341 del capítulo 11, hallar los valores propios y los vectores propios de f.

473. Las mismas preguntas que en el ejercicio precedente para el endomorfismo estudiado en el ejercicio 342 del capítulo 11.

474. Si E designa el espacio de los polinomios con coeficientes complejos de grado $\leq 2n$ se considera la aplicación f de E en $\mathbf{C}[X]$ definida por

$$P \rightarrow Q = f(P) = (X - a)(X - b)P' - [2nX - n(a + b) + c]P$$

donde a, b, c son números complejos tales que $c/(a - b)$ no sea un entero racional.

a) Demostrar que f es un endomorfismo de E.

b) Hallar valores propios y los vectores propios de f. Deducir que f es automorfismo de E.

c) Hallar el polinomio P tal que $f(P) = 1$. (Se utilizará una base de E formada por vectores propios.)

475. Se considera un espacio vectorial E complejo de dimensión finita $n + 1$ y tres endomorfismos u, v, h de E satisfaciendo las relaciones siguientes

$$\text{(I)} \qquad h \circ u - u \circ h = 2u, \qquad h \circ v - v \circ h = -2v, \qquad u \circ v - v \circ u = -h.$$

1. Sea α un valor propio de h; se considera un vector $x \in E$ tal que $h(x) = \alpha x$; demostrar que los vectores $y = u(x)$ y $z = v(x)$ verifican las relaciones

$$h(y) = (\alpha + 2)y, \qquad h(z) = (\alpha - 2)z$$

deducir que se tiene $h[u^k(x)] = (\alpha + 2k)u^k(x)$ para todo entero k positivo. Demostrar que existe un $k \geq 0$ tal que se tiene $u^k(x) = 0$.

2. Demostrar que existe un número complejo α y un vector $x \in E$ no nulo tales que

$$h(x) = \alpha x, \qquad u(x) = 0.$$

3. Si x y α son los mismos que en la pregunta 2, se pone

$$x_k = v^k(x)/k! \qquad (k = 0, 1, 2, \ldots).$$

Establecer las relaciones

$$h(x_k) = (\alpha - 2k)x_k, \qquad v(x_k) = (k+1)x_{k+1}, \qquad u(x_k) = (k-1-\alpha)x_{k-1}.$$

4. Se supone que fuera de E y de $\{0\}$ no hay en E ningún subespacio vectorial F tal que u, v, h apliquen F en F mismo. Demostrar que los vectores x_k de la pregunta 3 engendran E, a continuación que los x_k no nulos forman una base de E. Demostrar que de hecho los vectores

$$x_0, x_1, \ldots, x_n$$

forman una base de E, y calcular α en función de n.

5. Se toma por E el espacio vectorial formado por los polinomios de grado n a lo sumo de una variable t, con coeficientes complejos. Se considera los endomorfismos u, v, h de E definidos como sigue:

u aplica el polinomio $f(t)$ sobre el polinomio $-ntf(t) + t^2f'(t)$;
v aplica el polinomio $f(t)$ sobre el polinomio $f'(t)$;
h aplica el polinomio $f(t)$ sobre el polinomio $-nf(t) + 2tf'(t)$.

Demostrar que u, v, h satisfacen las relaciones (I) y a la condición enunciada al principio de la pregunta 4. Calcular en este caso los elementos x_k de la pregunta 3. (M.G.P.)

476. Sea f un endomorfismo de un espacio vectorial sobre K, (λ_h) $(1 \leq h \leq m)$ una familia de valores propios todos distintos de f. Se supone que existe una familia (k_h) $(1 \leq h \leq m)$ de enteros positivos no nulos y una familia (x_h) $(1 \leq h \leq m)$ de vectores no nulos de E tales que

$$(h = 1, 2, \ldots, m) \qquad (f - \lambda_h e)^{k_h}(x_h) = 0.$$

Demostrar que $x_1, \ldots, x_m$ son independientes.

477. Si f y g son dos endomorfismos del espacio vectorial E de dimensión n sobre **K**, teniendo cada uno n valores propios distintos 2 a 2 en **K**, demostrar que las propiedades son equivalentes:

a) $f \circ g = g \circ f$,

b) f y g tienen los mismos vectores propios.

478. Sean f y g dos endomorfismos permutables de un espacio vectorial E de dimensión finita.

a) Demostrar que todo subespacio de f es estable por g.

b) Demostrar por inducción sobre $n = \dim.$ E, que hay un vector propio $x \neq 0$ común a f y g.

479*. Sean f y g dos endomorfismos permutables de un espacio vectorial E de dimensión finita sobre un cuerpo K algebraicamente cerrado.

a) Demostrar que si f y g son diagonalizables existe una misma base (a_i) tal que $M(f, (a_i))$ y $M(g, (a_i))$ son diagonales.

b) Demostrar que si f y g son reducibles a la forma triangular, existe una misma base (a_i) tal que $M(f, (a_i))$ y $M(g, (a_i))$ son triangulares (utilizar el ejercicio precedente).

480. Sea E un espacio vectorial de dimensión n sobre el cuerpo K, $(e_1, e_2, \ldots, e_n)$ una base de E, f una aplicación lineal de E en sí mismo.
Sea x_1 un vector de E, se pone:

$$x_2 = f(x_1), \quad \ldots, \quad x_k = f(x_{k-1}) = f^{k-1}(x_1), \quad \text{etc.}$$

a) Demostrar que el subespacio F engendrado por los vectores x_k para $k = 1, 2, \ldots$ es estable por la aplicación f.

b) Sea p el mayor entero tal que $x_1, x_2, \ldots, x_p$ sean linealmente independientes. Demostrar que $x_1, x_2, \ldots, x_p$ forman una base de F. Si F coincide con todo E, se dirá en lo que sigue que x_1 *«engendra»* E. ¿Cuál es entonces el valor de p?

c) Se supone que la matriz A de f respecto a la base $(e_1, e_2, \ldots, e_n)$ es diagonal. Demostrar que, para que exista un vector x_1 *«engendrando»* E, es necesario y suficiente que los valores propios de f sean dos a dos distintos (se pondrá $x_1 = \Sigma\, \xi_i e_i$, y se buscará la condición para que los vectores $x_1, \ldots, x_n$ sean independientes).
Dar un ejemplo (para $n = 2$) de aplicación f (de matriz no reducible a la forma diagonal), teniendo dos valores propios iguales y tal que, sin embargo, existe un vector x_1 *«engendrando»* E.

d) Se supone que el vector x_1 *«engendra»* E; ¿cuál es la forma de la matriz A′ de f respecto a la base $(x_1, \ldots, x_n)$? Se designará por a_i el elemento situado en i-ésima fila y la última columna de esta matriz. Demostrar que el polinomio característico de A′ es

$$P(\lambda) = \det\,(f - \lambda e) = (-1)^n[\lambda^n - a_n\lambda^{n-1} - a_{n-1}\lambda^{n-2} - \ldots - a_1]$$

si e designa la aplicación idéntica de E en E. Demostrar que la matriz obtenida reemplazando en $P(\lambda)$ la variable λ por la matriz A′ de f es nula. (Se podrá demostrar que la aplicación lineal $f^n - a_n f^{n-1} - \ldots - a_2 f - a_1 e$ anula cada uno de los vectores de la base $(x_1, \ldots, x_n)$.

(M.G.P.)

481. Se considera el espacio vectorial E de dimensión finita n sobre el cuerpo **C** y los operadores lineales sobre E, f y g. Se designa por e la aplicación idéntica de E.

a) Se supone que $g \circ f$ tiene un valor propio nulo; demostrar que $f \circ g$ tiene también un valor propio nulo.

b) Dar la condición necesaria y suficiente referente al espectro de un operador lineal para que $e - f$ sea inversible.

c) Sea $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ el espectro del operador f. ¿Se puede elegir el número α complejo de modo tal que $e - \alpha f$ sea no inversible?

d) Si f y g son dos operadores lineales, se supone que $e - (f \circ g)$ es inversible. Demostrar que

$$[e - (g \circ f)] \circ [e + g \circ (e - f \circ g)^{-1} \circ f] = e$$

deducir que $e - (g \circ f)$ es inversible.

e) De lo precedente demostrar que, cualesquiera que sean los operadores lineales f y g, $f \circ g$ y $g \circ f$ tienen los mismos valores propios. **(M.G.P.)**

482. Se consideran dos sucesiones (u_n) y (v_n) de números complejos tales que para todo entero natural n

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = A \begin{pmatrix} u_{n-1} \\ v_{n-1} \end{pmatrix}.$$

Si A es una matriz de $M_2(\mathbf{C})$ independiente de n, sean s_1 y s_2 los valores propios de A. Calcular u_n y v_n en función de u_0, v_0, n y los elementos de A:

a) cuando $s_1 \neq s_2$, **b)** cuando $s_1 = s_2 = s$ (se deducirá primero el caso en que A es triangular inferior, buscando v'_n verificando $v_n = s^n v'_n$).

483. Se considera k sucesiones complejas $(u_{1,n})$, ..., $(u_{k,n})$ tales que todo entero natural n

$$\begin{pmatrix} u_{1,n+1} \\ u_{2,n+1} \\ \cdot \\ \cdot \\ \cdot \\ u_{k,n+1} \end{pmatrix} = A \begin{pmatrix} u_{1,n} \\ u_{2,n} \\ \cdot \\ \cdot \\ \cdot \\ u_{k,n} \end{pmatrix}.$$

Si A es una matriz de $M_k(\mathbf{C})$ independiente de n.
Calcular, para $h = 1, 2, ..., k$, $u_{h,n}$ en función de $u_{1,0}$, ..., $u_{k,0}$ de n y de los elementos de A cuando los valores propios de A son todos distintos.

484. Se consideran las sucesiones complejas u_n verificando para todo entero natural n (k, entero natural ≥ 1, fijo)

$$(1) \qquad u_{n+k} = a_0 u_n + a_1 u_{n+1} + \dots + a_{k-1} u_{n+k-1}$$

donde a_0, ..., a_{k-1} son números complejos dados.

a) Demostrar que estas sucesiones describen un subespacio vectorial de $\mathcal{F}(\mathbf{N}, \mathbf{C})$, espacio vectorial sobre $\mathbf{C}$ (V. cap. 7, ej. 149). ¿Cuál es su dimensión? (Se considerará las sucesiones $(e_{h,n})$ que verifican (1) y tales que $e_{h,l} = \delta_{hl}$ para $0 \leq h \leq k-1$).

b) Demostrar que el cálculo de u_n se reduce a hallar las k sucesiones definidas en el ejercicio 483. Se pondrá $u_{1,n} = u_n$, $u_{2,n} = u_{n+1}$, ..., $u_{k,n} = u_{n+k-1}$; ¿cuál es la matriz A correspondiente?

c) Calcular u_n para las sucesiones definidas en el ejercicio 149 del capítulo 7.

485. Si f es un endomorfismo de un espacio vectorial E de dimensión n sobre un cuerpo $\mathbf{K}$ algebraicamente cerrado. Se supone que el polinomio característico p_f es tal que $p_f = p_1 p_2$, con p_1 y p_2 dos polinomios *distintos* de $\mathbf{K}[X]$. Se pone

$$f_1 = p_1(f), \qquad f_2 = p_2(f), \qquad N_1 = \text{Ker } f_1, \qquad N_2 = \text{Ker } f_2.$$

a) Demostrar que N_1 y N_2 *son estables para* f.

b) Demostrar que $f_1(E) \subset N_2$, $f_2(E) \subset N_1$ (utilizar el teorema de Cayley-Hamilton).

c) Demostrar que $E = N_1 + N_2$ (se observará que existe q_1 y q_2 de $\mathbf{K}[X]$ tal que $p_1 q_1 + p_2 q_2 = 1$, se demostrará que todo x de E puede escribirse

$$x = f_1[q_1(f)(x)] + f_2[q_2(f)(x)]$$

y se aplicará el resultado de la pregunta **b**).
Demostrar en fin que $N_1 \cap N_2 = \{0\}$; luego que $E = N_1 \oplus N_2$.

d) Si g_i es la aplicación inducida por f en $N_i (i = 1, 2)$, demostrar que la matriz de f respecto a una base de E reunión de bases de N_1 y N_2 es de la forma

$$M = \begin{pmatrix} M_1 & O \\ O & M_2 \end{pmatrix}.$$

Deducir que $p_f = p_1 p_2 = p_{g_1} p_{g_2}$ (utilizar el ejercicio 262 del cap. 9).
Demostrar que existe $k \neq 0$ de $\mathbf{K}$ tal que $p_{g_1} = k_1 p_1$ (efectuar la división euclídea de $p_1(X)$ por $X - \lambda$, siendo λ un valor propio de g_1).

486* Sea f un endomorfismo de espacio vectorial E de dimensión n sobre **K** algebraicamente cerrado. Si el valor propio $\lambda_h (1 \leq h \leq m)$ es de orden k_h

$$(k_1 + k_2 + \dots + k_m = n)$$

se pone

$$(h = 1, \dots, m) \qquad N_h = \text{Ker } (f - \lambda_h e)^{k_h}.$$

a) Demostrar que

$$E = N_1 \oplus N_2 \oplus \dots \oplus N_m.$$

b) ¿Cuál es la forma de la matriz de f respecto a una base de E que sea reunión de bases de N_h? Deducir que dim $N_h = k_h$.

c) Demostrar que existen unas bases de E tales que M(f) sea un cuadro diagonal de matrices cuadradas $M_h (1 \leq h \leq m)$, siendo M_h triangular de orden k_h y con sus elementos diagonales todos iguales a λ_h. (Utilizar el ejercicio anterior).

487. Sea la matriz de $M_5(\mathbf{C})$

$$A = \begin{pmatrix} 1 & 1 & -1 & 2 & -1 \\ 2 & 0 & 1 & -4 & -1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & -3 & 3 & -1 \end{pmatrix}.$$

a) Calcular el polinomio característico de A. ¿Cuáles son sus raíces?

b) Hallar una matriz B semejante a A, cuadro diagonal de dos matrices triangulares (utilizar el ejercicio precedente).

488*. Si f es un endomorfismo del espacio vectorial E sobre el cuerpo **K**.

a) Demostrar que $e, f, f^2, \dots, f^{(n^2)}$ son linealmente dependientes. Deducir de ello la existencia de un polinomio unitario único $\omega_f \in \mathbf{K}[X]$ de grado mínimo k tal que $\omega_f(f) = 0$: se dirá que ω_f es el *polinomio minimal de f*.

b) Si $A = M(f, (a_i))$ demostrar que ω_f es también el polinomio unitario único de grado mínimo tal que $\omega_f(A) = 0$: se dirá también que ω_f es el polinomio minimal de A. Demostrar que dos matrices semejantes tienen igual polinomio minimal.

c) Volver a hallar la noción de polinomio minimal de f demostrando que el conjunto de los polinomios p de $\mathbf{K}[X]$ tales que $p(f) = 0$ es un ideal de $\mathbf{K}[X]$. (Utilizar § 186, ej. 5.)

d) Siendo λ un valor propio de f, demostrar que para todo entero $h \geq 2$, λ^h es valor propio de f. Deducir que todo valor propio de f es raíz del polinomio *minimal de f*.

e) Utilizando el teorema de CAYLEY-HAMILTON demostrar que el polinomio característico de f es un múltiplo de su polinomio minimal.

489*. Se toma las mismas notaciones que en el ejercicio anterior, se supone que **K** es algebraicamente cerrado y se pone

$$p_f = (-1)^n (X - \lambda_1)^{k_1} \dots (X - \lambda_m)^{k_m}$$

siendo $\lambda_1, \dots, \lambda_m$ distintos dos a dos, $k_1 + k_2 + \dots + k_m = n$.

a) Demostrar que

$$\omega_f = (X - \lambda_1)^{h_1} \dots (X - \lambda_m)^{h_m}$$

con $1 \leq h_i \leq k_i$ para $i = 1, 2, \dots, m$.

b) Demostrar que si f es diagonalizable, su polinomio minimal sólo tiene raíces simples.

c) Sea $M = (\mu_i^j)$ una matriz cuadrada tal que:

1. $j > i$ implica $\mu_i^j = 0$;
2. para todo i, $\mu_i^i = \lambda$;
3. existe i y j tales que $j < i$ y $\mu_i^j \neq 0$;

demostrar que λ no puede ser raíz simple del polinomio minimal de M.
Deducir que si el polinomio minimal de f sólo tiene raíces simples f es diagonalizable.

490*. Sea f un endomorfismo de un espacio vectorial de dimensión finita sobre **K**. Se considera el subanillo $\mathbf{K}[f]$ del anillo $\mathcal{L}(E)$. Se recuerda que $\mathbf{K}[f]$ está engendrado por $K \cup \{f\}$. Sea finalmente ω_f el polinomio minimal de f.

a) Demostrar que los divisores de cero del anillo $\mathbf{K}[f]$ son los endomorfismos $p(f)$ donde p es un polinomio de K[X] no primo con ω_f y de grado estrictamente inferior al de ω_f.

b) Suponiendo que p sea primo con ω_f, demostrar que $p(f)$ es un automorfismo de E.

491. Siendo f un endomorfismo de E espacio vectorial de dimensión finita sobre **K** se designa por φ el homomorfismo de $\mathbf{K}[X]$ en $\mathbf{K}[f]$ (ver ej. 490) definido por $\varphi(p) = p(f)$.
Demostrar que el núcleo de φ es el ideal (ω_f) engendrado por el polinomio minimal de f. Deducir de lo anterior que $K[f]$ es isomorfo a $\mathbf{K}[X]/(\omega_f)$.

492. Si f es el endomorfismo de $\mathbf{R}^3$ tal que en relación a la base canónica de $\mathbf{R}^3$

$$M(f) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

hallar los divisores de cero del anillo $\mathbf{R}[f]$ (V. ej. 490). Precisar los rangos de los endomorfismos hallados. (Se observará que $f^3 = e$.)

493.* Sea f un endomorfismo del espacio vectorial E de dimensión finita sobre **K** algebraicamente cerrado. Sea ω_f el polinomio minimal de f que se supone de grado k. Se designa por $Q(X, Y)$ el polinomio de $\mathbf{K}[X, Y]$ definido por

$$\omega_f(X) = (X - Y)Q(X, Y) + \omega_f(Y)$$

a) Demostrar que el polinomio Q es simétrico y de grado $k-1$ a lo sumo

b) Demostrar que si $\mu \in \mathbf{K}$ no es valor propio de f,

$$(f - \mu e)^{-1} = -\frac{Q(f, \mu)}{\omega_f(\mu)}.$$

c) Siendo h_i el orden de multiplicidad de λ_i en el polinomio minimal de f, demostrar que existen los polinomios $Q_{ij} \in \mathbf{K}[X]$ tales que

$$(f - \mu e)^{-1} = \sum_{i=1}^{h} \sum_{j=1}^{h_i} \frac{Q_{ij}(f)}{(\mu - \lambda_i)^j}.$$

siendo h el número de raíces distintas de ω_f.

494*. Sea E un espacio vectorial de dimensión n sobre K, se considera un endomorfismo nilpotente de E, es decir f tal que existe $p \geq 1$ tal que $f^{p-1} \neq 0$, $f^p = 0$.

Se pone

$$(k = 0, 1, \ldots, p) \qquad \mathrm{N}_k = \mathrm{K}er\,(f^k).$$

a) Demostrar que la sucesión de subespacios vectoriales de E

$$\{0\} = \mathrm{N}_0, \mathrm{N}_1, \ldots, \mathrm{N}_{p-1}, \mathrm{N}_p = \mathrm{E}$$

es estrictamente creciente. Deducir que $p \geq n$ (V. cap. 7, ej. 160).

b) Demostrar que, para todo $k > 0$, $f(\mathrm{N}_{k+1}) \subset \mathrm{N}_k$.

c) Sea F un subespacio de E tal que $\mathrm{F} \cap \mathrm{N}_k = \{0\}$ para $1 \leq k \leq p-1$; demostrar que $f(\mathrm{F}) \cap \mathrm{N}_{k-1} = \{0\}$ y que la restricción de f a F es inyectiva.

d) Demostrar que hay una sucesión $\mathrm{F}_1, \mathrm{F}_2, \ldots, \mathrm{F}_p$ de subespacios de E tales que, para $1 \leq k \leq p$, $\mathrm{N}_k = \mathrm{F}_k \oplus \mathrm{N}_{k-1}$, donde f aplica F_k inyectivamente en F_{k-1} $(k \geq 2)$. (Tomar como F_p un suplemento de N_{p-1} en $\mathrm{E} = \mathrm{N}_p$, a continuación F_{p-1} un suplementario de N_{p-2} en N_{p-1}, etc.). Demostrar que $\mathrm{F}_1 = \mathrm{N}_1$ y que E es suma directa de $\mathrm{F}_1, \mathrm{F}_2, \ldots, \mathrm{F}_p$.

e) Sea $a_{p,1}, a_{p,2}, \ldots, a_{p,n_p}$ una base de F_p, demostrar que

$$a_{p-1,1} = f(a_{p,1}), \ \ldots, \ a_{p-1,n_p} = f(a_{p,n_p}), \ a_{p-1,n_p+1}, \ \ldots, \ a_{p-1,n_{p-1}}$$

es una base de F_{p-1}. Formar igualmente bases de

$$\mathrm{F}_{p-2}, \ldots, \mathrm{F}_1 = \mathrm{N}_1 = \mathrm{Ker}\,(f).$$

La reunión de estas bases es una base de E descrita por $a_{k,l}$; se ordena estos elementos de manera que $a_{k,l}$ sea anterior a $a_{k',l'}$ si y solamente si

$$k < k' \quad \text{o} \quad k \leq k', \ l < l' \quad \text{sea} \quad (b_i) \quad (1 \leq i \leq n)$$

la base de E así obtenida. Demostrar que para todo i de $[1, n]$ $f(b_i) = 0$ o bien $f(b_i) = b_{i-1}$.
De ello deducir que $(\beta_i^j) = \mathrm{M}(f, (b_i))$ tiene todos sus elementos nulos excepto ciertos elementos β_i^{i-1} iguales a 1.

495*. Se llama *matriz de Jordan* de orden m, relativa a $\lambda \in \mathrm{K}$, toda matriz de $\mathrm{M}_m(\mathbf{K})$, representada $\mathrm{U}_{m,\lambda} = (\alpha_i^j)$ tal que

$$(i = 1, 2, \ldots, m) \ \alpha_i^i = \lambda: \qquad (i = 2, 3, \ldots, m) \ \alpha_i^{i-1} = 1$$

siendo nulos todos los demás elementos.

a) Demostrar que toda matriz de un endomorfismo nilpotente de un espacio vectorial de dimensión finita es semejante a un cuadro diagonal de matriz de JORDAN relativo a $\lambda = 0$. (Utilizar el ejercicio anterior y observar que $\mathrm{U}_{1,0} = (0)$).

b) Demostrar que toda matriz A de un endomorfismo f de un espacio vectorial de dimensión finita sobre un cuerpo $\mathbf{K}$ algebraicamente cerrado es semejante a un cuadro diagonal de matrices de JORDAN relativas a $\lambda_1, \ldots, \lambda_m$ valores propios de f. Este cuadro diagonal se llama *reducida* de JORDAN de A (utilizar el ejercicio 486). Con las notaciones de este ejercicio si f_h es la restricción de f a N_h se demostrará que $f_h = \lambda_h e + g_h$ para $1 \leq h \leq m$, siendo g_h un endomorfismo de N_h nilpotente de índice k_h orden de multiplicidad de λ_h en el polinomio característico de f.

496*. a) ¿Cuál es el polinomio minimal de la matriz de JORDAN $U_{m,\alpha}$?

b) Demostrar que si A es semejante a un cuadro diagonal de matrices de JORDAN $U_{m_i,\alpha}$, el polinomio minimal de A es el m. c. m. de los polinomios $(X - \alpha_i)^{m_i}$.

c) Demostrar que un endomorfismo f de un espacio vectorial de dimensión finita sobre un cuerpo algebraicamente cerrado es diagonalizable si y solamente si su polinomio minimal tiene solamente raíces simples. (Utilizar una reducida de JORDAN de f.)

Hallar las reducidas de Jordan de las matrices siguientes (se observará que el polinomio característico de cada una de ellas tiene raíces múltiples).

497.

$$\begin{pmatrix} 0 & 3 & 3 \\ -1 & 8 & 6 \\ 2 & -14 & -10 \end{pmatrix}.$$

498.

$$\begin{pmatrix} 3 & 2 & -3 \\ 4 & 10 & -12 \\ 3 & 6 & -7 \end{pmatrix}.$$

499. Matriz del ejercicio 462 tomando $a = b = c = d = 1$.

500. Matriz del ejercicio 487.

15

FORMAS BILINEALES SIMETRICAS Y FORMAS HERMITIANAS

En las tres primeras secciones estudiaremos las formas bilineales y las formas cuadráticas sobre E, espacio vectorial sobre K, *siendo* K *un cuerpo de característica* $\neq 2$. En la sección IV estudiaremos las propiedades particulares de las formas reales, y, en particular, el *espacio vectorial euclídeo real de dimensión n.*

Finalmente en la sección V estudiaremos las formas sesquilineales y hermitianas(*) sobre E, espacio vectorial sobre **C**, y, en particular, el *espacio vectorial hermitiano complejo de dimensión n.*

I. Definición y primeras propiedades de las formas bilineales y de las formas cuadráticas

221. Formas bilineales definidas sobre E $\times$ F

a) Espacio vectorial $\mathcal{L}_2$(E, F; K)

Sean E y F dos espacios vectoriales sobre K, hemos definido en el § 164 las formas bilineales sobre E $\times$ F (def. 2), y hemos visto que el conjunto de estas formas, $\mathcal{L}_2$(E, F; K) provisto de la adición y de la multiplicación para un escalar, tiene una estructura de espacio vectorial sobre K (t. 1).

(*) *N. del T.* — Aunque algunos autores denominan a estas formas «hermíticas», hemos preferido —por considerarlo más castellano— la denominación «hermitiana», si bien convendría para castellanizarla completamente denominarla «hermiciana»; por razones fonéticas no lo hemos hecho.

Supongamos ahora que E y F son de dimensión finita, y designemos por $(a_i)(1 \leq i \leq m)$ y $(b_j)(1 \leq j \leq n)$ una base de E y una base de F. Pongamos

$$x = \sum_{i=1}^{m} x^i a_i \qquad y = \sum_{j=1}^{n} y^j b_j$$

para toda familia bilineal f definida sobre $E \times F$ tendremos, gracias a las dos propiedades de bilinealidad

$$f(x, y) = f\left(\sum_i x^i a_i, \quad \sum_j y^j b_j\right) = \sum_i \sum_j x^i y^j f(a_i, b_j)$$

de donde poniendo $\alpha_{ij} = f(a_i, b_j)$

$$(1) \qquad f(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} x^i y^j$$

f permite, pues, determinar mn escalares α_{ij}, verificando (1). Recíprocamente dado mn escalares α_{ij}, siendo escogidas unas bases en E y F; la fórmula (1) define una aplicación de $E \times F$ en K, que es manifiestamente bilineal. Consideremos, en particular, las mn formas bilineales φ^{ij} definidas por

$$(i = 1, \ldots, m; \qquad j = 1, \ldots, n) \qquad \varphi^{ij}(a_{i'}, b_{j'}) = \delta^i_{i'} \delta^j_{j'}$$

donde $\delta^i_{i'}$ y $\delta^j_{j'}$ son los símbolos de KRONECKER, tenemos, pues,

$$\varphi^{ij}(x, y) = x^i y^j$$

luego para toda forma bilineal f de $\mathfrak{L}_2(E, F; K)$

$$(\forall x \in E)(\forall y \in F) \qquad f(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} \varphi^{ij}(x, y)$$

es decir,

$$f = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} \varphi^{ij}$$

luego las mn formas φ^{ij} engendran $\mathfrak{L}_2(E, F; K)$; estas formas son visiblemente independientes; en efecto,

$$l = \sum_{i=1}^{m} \sum_{j=1}^{n} \lambda_{ij} \varphi^{ij} = 0$$

implica para todo i' de $[1, m]$ y todo j' de $[1, n]$, $l(a_{i'}, b_{j'}) = \lambda_{i'j'} = 0$, luego:

TEOREMA 1. — *Si* E *y* F *son dos espacios vectoriales de dimensiones respectivas* m *y* n *sobre* K, $\mathfrak{L}_2(E, F; K)$ *es un espacio vectorial de dimensión* mn *sobre* K.

b) Matrices asociadas a una forma bilineal

Elegidas unas bases en E y F de dimensión finita, podemos hacer corresponder a todo elemento f de $\mathfrak{L}_2(\mathrm{E}, \mathrm{F}; \mathrm{K})$ la matriz $\mathrm{A} = (\alpha_{ij})$ (i, índice de columna; j, índice de fila) que es un elemento de $\mathbf{M}_{\mathrm{K}}(m, n)$. Las fórmulas $f(a_i, b_j) = \alpha_{ij}$, muestran que, *escogidas las bases*, la aplicación de $\mathfrak{L}_2(\mathrm{E}, \mathrm{F}; \mathrm{K})$ en $\mathbf{M}_{\mathrm{K}}(m, n)$ definida por $f \to \mathrm{A}$, es una biyección; se ve inmediatamente que esta aplicación es un *isomorfismo* de los espacios vectoriales $\mathfrak{L}_2(\mathrm{E}, \mathrm{F}; \mathrm{K})$ y $\mathbf{M}_{\mathrm{K}}(m, n)$. *Se dice que A es la* matriz asociada a la forma bilineal f *relativamente a las bases* (a_i) *de* E *y* (b_j) de F.

Según (1) tenemos (siendo K conmutativo)

$$f(x, y) = \sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_{ij}x^i y^j = \sum_{j=1}^{n} y^j(\alpha_{1j}x^1 + \ldots + \alpha_{mj}x^m)$$

es decir,

$$f(x, y) = (y^1 \ldots y^n)\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{m1} \\ \vdots & & \vdots \\ \alpha_{1n} & \cdots & \alpha_{mn} \end{pmatrix}\begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix}$$

o, también, poniendo $\mathrm{X} = \mathrm{M}(x, (a_i))$, $\mathrm{Y} = \mathrm{M}(y, (b_j))$ (ver § 154, ejemplo 1).

(2) $$f(x, y) = {}^t\mathrm{YAX}$$

la matriz ${}^t\mathrm{YAX}$ siendo una matriz (1, 1) es igual a su transpuesta, luego

(2′) $$f(x, y) = {}^t\mathrm{X}\,{}^t\mathrm{AY}$$

fórmula que se podría demostrar directamente a partir de (1), de donde:

TEOREMA 2. — *Elegidas unas bases* (a_i) $(1 \leq i \leq m)$, (b_j) $(1 \leq j \leq n)$, *respectivamente, en* E *y* F, *a toda forma bilineal* f *definida sobre* $\mathrm{E} \times \mathrm{F}$, *se puede asociar la matriz* $\mathrm{A} = (\alpha_{ij})$ *(i índice de columna, j índice de fila) de tipo* (m, n) *definida por* $\alpha_{ij} = f(a_i, b_j)$.

La aplicación definida por $f \to \mathrm{A}$ *es un isomorfismo de* $\mathfrak{L}_2(\mathrm{E}, \mathrm{F}, \mathrm{K})$ *sobre* $\mathrm{M}_{\mathrm{K}}(m, n)$.

Si X *e* Y *representan las matrices unicolumnas de las coordenadas de* x *e* y *relativamente a las bases* (a_i) *y* (b_j), *respectivamente, se tiene*

$$f(x, y) = {}^t\mathrm{YAX} = {}^t\mathrm{X}\,{}^t\mathrm{AY}.$$

Supongamos que cambiamos de base en E y en F, sea P la matriz de paso de la base (a_i) a la base (a'_i) de E y Q, la matriz de pasaje de la base (b_j) a la (b'_j) de F. Poniendo $\mathrm{X}' = \mathrm{M}(x, (a'_i))$, $\mathrm{Y}' = \mathrm{M}(y, (b'_j))$, tendremos, gracias a la relación (2), a las relaciones $\mathrm{X} = \mathrm{PX}'$, $\mathrm{Y} = \mathrm{QY}'$ (ver § 160) y a las propiedades de la multiplicación de matrices

$$f(x, y) = {}^t\mathrm{YAX} = {}^t(\mathrm{QY}')\mathrm{A}(\mathrm{PX}') = {}^t\mathrm{Y}'({}^t\mathrm{QAP})\mathrm{X}'$$

vemos entonces que relativamente a las nuevas bases, a f está asociada la matriz

(3) $$\mathrm{A}' = {}^t\mathrm{QAP}.$$

Se observará la *analogía* y la *diferencia* entre la fórmula (3) anterior y la fórmula $A' = Q^{-1}AP$ del § 161, correspondiente a las matrices asociadas a una *aplicación lineal* de E en F.

EJERCICIOS

1. $E = F = K^n$, E está expresado en su base canónica, cuál es la matriz asociada a la forma bilineal definida sobre $E \times E$ por

$$f(x, y) = x^1y^1 + \ldots + x^ny^n.$$

2. Demostrar que en $M_K(m, n)$ la relación: «existe P de $GL_m(K)$ y Q de $GL_n(K)$, tales que $A' = {}^tQAP$» es una relación de equivalencia.

222. Formas bilineales definidas sobre $E \times E$

a) Espacio vectorial $\mathcal{L}_2(E; K)$

Si $E = F$ definimos una forma bilineal f sobre $E \times E$, su conjunto describe un espacio vectorial representado $\mathcal{L}_2(E; K)$ (ver § 164). Para simplificar diremos que f es una *forma bilineal sobre* E.

Si E es de dimensión n, $\mathcal{L}_2(E; K)$ es de dimensión n^2; si (a_i) es una base de E ($i \in [1, n] = I$) a toda forma bilineal f sobre E, podemos asociar una matriz cuadrada $A = (\alpha_{ij})$ ($(i, j) \in I \times I$); existe un isomorfismo entre los espacios vectoriales $\mathcal{L}_2(E; K)$ y $\mathbf{M}_n(K)$. Si (a'_i) es otra base de E, relativamente a esta nueva base la matriz asociada a f es

$$(3') \qquad A' = {}^tPAP.$$

Siendo F igual a E, podemos definir en $\mathcal{L}_2(E; K)$ las *formas bilineales simétricas,* que están definidas por

$$(\forall (x, y) \in E^2) \qquad f(x, y) = f(y, x).$$

Las fórmulas $\alpha_{ij} = f(a_i, a_j)$ muestran que en este caso la *matriz asociada a f* relativamente a cualquier base de E, supuesta de dimensión finita, es *simétrica* si y solamente si f es simétrica.

Podemos igualmente definir en $\mathcal{L}_2(E; K)$ las *formas bilineales antisimétricas,* que están definidas por

$$(\forall (x, y) \in E^2) \qquad f(y, x) = -f(x, y)$$

siendo K de característica diferente de 2, se demuestra fácilmente (ver § 165, ejercicio) que f, forma bilineal sobre E es *antisimétrica* si y sólo si es *alternada,* es decir, si

$$(\forall x \in E) \qquad f(x, x) = 0.$$

Hemos visto en el § 165 que el conjunto de las formas bilineales alternadas sobre E es un subespacio vectorial $\mathcal{A}_2(E; K)$ de $\mathcal{L}_2(E; K)$ (hacer $n = 2$ en el teorema 3 del § 165). Está igualmente claro que el conjunto de las formas bilineales simétricas sobre E es un subespacio $\mathcal{S}_2(E; K)$ de $\mathcal{L}_2(E; K)$. Estos dos subespacios son suplementarios (ver ejercicio 3, más abajo).

EJEMPLOS Y EJERCICIOS

1. El ejercicio 1 del § 221, da un ejemplo de forma bilineal simétrica. Si $n = 3$ y $K = \mathbf{R}$, $f(x, y)$ no es otro que el producto escalar clásico de los vectores x e y.

2. Siendo E el espacio vectorial de las funciones reales continuas sobre $[\alpha, \beta]$, la aplicación f de E^2 en $\mathbf{R}$ definida por

$$(x, y) \to f(x, y) = \int_{\alpha}^{\beta} x(t)y(t)dt$$

es una forma bilineal simétrica sobre ($t \to x(t)$ y $t \to y(t)$ son los elementos de E).

3. S (resp. A) designa el conjunto de las formas bilineales sobre E simétricas (resp. *antisimétricas*), demostrar que $\mathfrak{L}_2(E, K) = S \oplus A$ (se recuerda que K es de característica $\neq 2$); este resultado, ¿también se verifica si K es de característica 2?

4. Se supone dim $E = n$, ¿cuáles son las dimensiones de S y de A? (ver § 156, ej. 1).

b) Isomorfismo entre $\mathfrak{L}_2(E; \mathbf{K})$ y $\mathfrak{L}(E, E^*)$

A toda forma bilineal f sobre E, podemos asociar, perteneciendo x a E, la aplicación parcial $f(x, *)$ que representaremos f_x; por definición de la bilinealidad (ver § 164) f_x es una aplicación lineal de E en K, luego un elemento del dual E^* de E. Además, según la definición de la forma bilineal canónica definida sobre $E \times E^*$ (ver § 149), para todo par (x, y) de $E \times E$ y todo elemento f de $\mathfrak{L}_2(E; K)$, tenemos

$$\langle y, f_x \rangle = f_x(y) = f(x, y).$$

Designemos por φ la aplicación de E en E^* definida por

$$\varphi: \ E \to E^*, \qquad x \to \varphi(x) = f_x$$

es lineal; en efecto, cualesquiera que sean x_1, x_2, y de E y λ de K

$$\begin{cases} f_{x_1+x_2}(y) = f(x_1 + x_2, y) = f(x_1, y) + f(x_2, y) = f_{x_1}(y) + f_{x_2}(y) \\ f_{\lambda x}(y) = f(\lambda x, y) = \lambda f(x, y) = \lambda f_x(y) \end{cases}$$

luego

$$\begin{cases} f_{x_1+x_2} = f_{x_1} + f_{x_2} \\ f_{\lambda x} = \lambda f_x \end{cases} \Leftrightarrow \begin{cases} \varphi(x_1 + x_2) = \varphi(x_1) + \varphi(x_2) \\ \varphi(\lambda x) = \lambda\varphi(x). \end{cases}$$

Entonces a todo elemento f de $\mathfrak{L}_2(E; K)$, asociamos un elemento φ de $\mathfrak{L}(E, E^*)$; designemos por φ la aplicación así definida

$$\theta: \ \mathfrak{L}_2(E; K) \to \mathfrak{L}(E, E^*), \qquad f \to \theta(f) = \varphi$$

vamos a demostrar que θ es un *isomorfismo de espacios vectoriales*.

Primero θ es *lineal*; cualquiera que sean f' y f'' de $\mathfrak{L}_2(E; K)$, pongamos $f = f' + f''$ y

$$\theta(f') = \varphi', \qquad \theta(f'') = \varphi'', \qquad \theta(f) = \varphi$$

para todo x y todo y de E, tenemos

$$[\varphi(x)]\,(y) = f_x(y) = f(x, y) = (f' + f'')\,(x, y) = f'(x, y) + f''(x, y)$$
$$= f'_x(y) + f''_x(y) = [\varphi'(x)]\,(y) + [\varphi''(x)]\,(y)$$

de donde, para todo x de E

$$\varphi(x) = \varphi'(x) + \varphi''(x)$$

es decir,

$$\varphi = \varphi' + \varphi'' \Leftrightarrow \theta(f' + f'') = \theta(f') + \theta(f'')$$

se demostrará igualmente que para todo f de $\mathcal{L}_2(E;\ K)$ y todo λ de K

$$\theta(\lambda f) = \lambda\theta(f).$$

Seguidamente θ es *inyectiva*: Pongamos $\varphi = \theta(f)$, $\varphi' = \theta(f')$ y supongamos $\varphi = \varphi'$; para todo x de E

$$\varphi(x) = \varphi'(x) \Leftrightarrow f_x = f'_x$$

luego para todo x y todo y de E

$$f_x(y) = f'_x(y) \Leftrightarrow f(x,\ y) = f'(x,\ y)$$

es decir, $f = f'$.

Demostremos, en fin, que θ es *suprayectiva*. Sea, en efecto, φ una aplicación lineal cualquiera de E en E*, luego para todo x de E, es una forma lineal definida sobre E; consideremos la aplicación f de E × E en K definida por

$$(x,\ y) \to f(x,\ y) = [\varphi(x)]\,(y)$$

f es bilineal: es lineal en x, pues φ es lineal y lineal en y puesto que $\varphi(x)$ es lineal. Entonces cualquiera que sea φ de $\mathcal{L}(E,\ E^*)$ existe f de $\mathcal{L}_2(E;\ K)$ tal que $\theta(f) = \varphi$. Por lo tanto:

TEOREMA 3. — *Siendo f una aplicación bilineal sobre* E, *definida por* $(x,\ y) \to f(x,\ y)$, *la aplicación de* E *en* E*, *definida por* $\varphi(x) = f_x$, *es lineal, y la aplicación* θ *de* $\mathcal{L}_2(E,\ K)$ *en* $\mathcal{L}(E,\ E^*)$, *definida por* $\theta(f) = \varphi$, *es un isomorfismo de espacios vectoriales.*

Se definiría igualmente f_y, que es también un elemento del dual E* de E, designado por ψ la aplicación de E en E* definida por $\psi(y) = f_y$, se obtendrá para todo par $(x,\ y)$ de elementos de E

$$\langle x,\ f_y \rangle = \langle x,\ \psi(y) \rangle = f(x,\ y).$$

De donde

$$\langle y,\ \varphi(x) \rangle = \langle x,\ \psi(y) \rangle = f(x,\ y)$$

en particular si f es *simétrica*, tendremos para todo x y para todo y

$$\langle y,\ \varphi(x) \rangle = \langle y,\ \psi(x) \rangle$$

es decir, $\varphi = \psi$.

COROLARIO. — *Si f es una forma bilineal simétrica sobre* E, *las dos aplicaciones lineales φ y ψ de* E *en* E*, *definidas por*

$$x \to \varphi(x) = f_x, \qquad y \to \psi(y) = f_y,$$

si f_x y f_y son las aplicaciones parciales asociadas a f, son idénticas.

OBSERVACION

El teorema 3 es válido en casos mucho más generales (ver § 164, ej. 1).

c) Caso en que E es de dimensión finita

Sea (a^{*i}) $(1 \leq i \leq n)$ una base de E, designemos por (α_{ij}) la matriz cuadrada de orden n asociada a la forma bilineal f definida sobre E × E, tenemos

$$(i,\ j = 1,\ \dots,\ n) \qquad \alpha_{ij} = f(a_i,\ a_j).$$

Sea (a^{*i}) $(1 \leq i \leq n)$ la base de E* dual de (a_i) (ver § 150), y designemos por (β_{ij}) la matriz

$$(\beta_{ij}) = \mathrm{M}(\varphi,\ (a_i),\ (a^{*j}))$$

es decir,

$$(i = 1,\ \dots,\ n) \qquad \varphi(a_i) = f_{a_i} = \sum_{k=1}^{n} \beta_{ik} a^{*k}$$

tendremos

$$\alpha_{ij} = f(a_i\ \ a_j) = f_{a_i}(a_j) = \sum_{k=1}^{n} \beta_{ik} a^{*k}(a_j) = \sum_{k=1}^{n} \beta_{ik} \langle a_j,\ a^{*k} \rangle = \sum_{k=1}^{n} \beta_{ik} \delta_j^k = \beta_{ij}.$$

Teorema 4. — E *de dimensión finita viene referido a una base* (a_i) *y* E* *en la base dual* (a^{*i}), *la matriz asociada a la forma bilineal* f, *relativamente a la base* (a_i) *es igual a la matriz asociada a* φ *relativamente a las bases* (a_i) *y* (a^{*i}).

Al determinante de la matriz (α_{ij}) se le llama el "*discriminante*" *de la forma bilineal* f *relativamente a la base* (a^i). Este determinante depende de la base escogida; en efecto, det A y det $({}^t\mathrm{PAP})$ son, en general, distintos; pero **P** al ser inversible

$$\det \mathrm{A} \neq 0 \Leftrightarrow \det\ ({}^t\mathrm{PAP}) \neq 0.$$

EJERCICIOS

5. Siendo E de dimensión finita, se puede identificar E** y E, una vez hecha esta identificación, demostrar que las dos aplicaciones φ y ψ consideradas anteriormente en el párrafo *b*) son transpuestas una de otra.

6. Siendo E de dimensión finita, utilizar el teorema 4 para demostrar en este caso que $\mathrm{L}_2(\mathrm{E},\ \mathrm{K})$ y $\mathcal{L}(\mathrm{E},\ \mathrm{E}^*)$ son isomorfos.

7. Demostrar, utilizando el teorema 4, que si P es la matriz de cambio de (a_i) a (a'_i), ${}^t\mathrm{P}$ es la matriz de cambio de (a^{*i}) a (a'^{*i}).

223. Formas bilineales simétricas y formas cuadráticas

Definición 1. — *Siendo* f *una forma bilineal cualquiera sobre* E, *se llama* forma cuadrática *asociada a* f *la aplicación* q *de* E *en* K *definida por*

$$(\forall\, x \in \mathrm{E}) \qquad q(x) = f(x,\ x).$$

Supongamos E de dimensión finita referido a una base (a_i), tendremos (§ 221, fórmula 1)

$$q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} x^i x^j$$

es decir, si consideramos q como una aplicación de K^n en K, q es una *función polinomia homogénea de grado 2*: La palabra «*forma*» al ser sinónima de «*polinomio homogéneo*», q se llama naturalmente forma cuadrática, la definición dada anteriormente generaliza esta definición a E de cualquiera dimensión.

Se ve inmediatamentae que el *conjunto* $\mathcal{Q}(E)$ *de las formas cuadráticas definidas sobre* E, espacio vectorial sobre K, provisto de la adición y de la multiplicación por un escalar es un espacio vectorial sobre K. Está igualmente claro que la aplicación de $\mathcal{L}_2(E;\ K)$ en $\mathcal{Q}(E)$ definida por $f \rightarrow q$ es un homomorfismo de espacios vectoriales. Pero este homomorfismo suprayectivo por definición no es inyectivo, en efecto, si a es una *forma bilineal* alternada no nula (ver § 165), se tendrá, por tanto, para todo x

$$f(x,\ x) = q(x), \qquad (f + a)(x,\ x) = f(x,\ x) + a(x,\ x) = q(x)$$

por tanto a las dos formas bilineales distintas f y $f + a$ está asociada la misma forma cuadrática q.

Pero supongamos f simétrica, siendo q la forma cuadrática asociada a f, tendremos, gracias a la bilinealidad y la simetría de f, con la *característica de* K *distinta de* 2

$$(\forall (x,\ y) \in E \times E) \qquad q(x + y) = f(x + y,\ x + y) = f(x,\ x) + f(x,\ y) + f(y,\ x) + f(y,\ y) = q(x) + q(y) + 2f(x,\ y)$$

es decir,

$$f(x,\ y) = \frac{1}{2}(q(x + y) - q(x) - q(y))$$

fórmula que muestra que si la misma forma cuadrática q está asociada a dos formas bilineales simétricas f y g, se tiene $f = g$, de donde:

TEOREMA 5. — *La aplicación que, a toda* forma bilineal simétrica f *sobre* E, *asocia la* forma cuadrática q definida por

$$(1) \qquad (\forall x \in E) \qquad q(x) = f(x,\ x)$$

es biyectiva.

La forma bilineal simétrica f, así asociada, a q es tal que

$$(2) \qquad (\forall (x,\ y) \in E \times E) \qquad f(x,\ y) = \frac{1}{2}(q(x + y) - q(x) - q(y))$$

se dice que f es la forma polar *de q.*

Las fórmulas (1) y (2) anteriores permiten enunciar los dos resultados siguientes:

COROLARIO 1. — *Si q es la forma cuadrática asociada a f forma bilineal simétrica sobre* E, *las relaciones $q = 0$ y $f = 0$ son equivalentes.*

COROLARIO 2. — *Si q es la forma cuadrática asociada a la forma bilineal simétrica f, para todo endomorfismo u de* E, *se tiene*

$$[(\forall x \in E) \quad q[u(x)] = q(x)] \Leftrightarrow [\forall (x, y) \in E^2 \quad f[u(x), u(y)] = f(x, y)]$$

se dice entonces que u conserva *f* (*o q*) (ver § 228).

Si E es de dimensión finita, la matriz $(\alpha^i_j) = A$ asociada a la forma bilineal simétrica f relativamente a una base (a^i) de E es aún llamada *matriz asociada* a q (relativamente a (a^i)), a su determinante se le llama también *discriminante* de q. La matriz (α_{ij}), siendo simétrica, y el cuerpo K (de característica $\neq 2$) conmutativo, tendremos

$$q(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij}x^i x^j = \sum_{i=1}^{n} \alpha_{ii}(\alpha^i)^2 + 2 \sum_{1 \leq i < j \leq n} \alpha_{ij}x^i x^j.$$

Se dice que $\alpha_{ii}(x^i)^2$ es un término "*cuadrado*" y $2\alpha^i_j x^i x^j$ un término "*rectangular*". La palabra "*cuadrado*" es un abuso de lenguaje, pues $\alpha_{ii}(x^i)^2$ no es un cuadrado en K más si α^i_i lo es: esto será siempre el caso en **C**, pero no en **R**.

Finalmente, se tiene igualmente con las notaciones del § 221, *b*)

$$q(x) = {}^t XAX.$$

EJEMPLOS Y EJERCICIOS

1. Si $E = K^n$, q definido por

$$q(x) = (x^1)^2 + \ldots + (x^n)^2$$

está asociada a la forma f definida por

$$f(x, y) = x^1y^1 + \ldots + x^ny^n$$

(ver § 221, ej. 1 y § 222, ej. 1).

2. Sea E el espacio vectorial de las funciones reales continuas sobre $[\alpha, \beta]$, q definido por

$$q(x) = \int_\alpha^\beta [x(t)]^2 dt$$

es una forma cuadrática definida sobre E (ver § 222, ej. 2).

3. Siendo u el homomorfismo de $\mathcal{L}_2(E, K)$ en $\mathcal{Q}(E)$ definido por $u(f) = q$, determinar el núcleo de u.

4. Tenemos E de dimensión n sobre K, referido a una base (a_i) se pone $x = \sum_{i=1}^{n} x^i a_i$, demostrar que si la aplicación q de E en K definido por $x \to q(x^1, \ldots, x^n)$ es tal que q sea un polinomio homogéneo de grado 2 en $x^1, \ldots, x^n$, lo mismo se verifica para cualquier otra base de E. Deducir que q es una forma cuadrática sobre E, y que toda forma cuadrática sobre E puede estar definida de esta forma.

5. Se toman las notaciones del ejercicio anterior, demostrar que la forma polar de q, expresada

$$(x^1, \ldots, x^n, y^1, \ldots, y^n) \to f(x^1, \ldots, x^n, y^1, \ldots, y^n)$$

está dada por la fórmula

$$f(x^1, \ldots, x^n, y^1, \ldots, y^n) = \frac{1}{2}\sum_{i=1}^{n} x^i q'_{x^i}(y^1, \ldots, y^n)$$

q'_{x^i} es la función derivada de q relativamente a x_i.

Demostrar en particular que

$$q(x) = (x^i)^2 \Rightarrow f(x, y) = x^i y^i$$
$$q(x) = 2x^i y^j \Rightarrow f(x, y) = x^i y^j + x^j y^i$$

estas dos reglas son conocidas con el nombre *«regla del desdoblamiento de variables»*.

6. Tomando las mismas notaciones que en los dos ejercicios anteriores, demostrar que el discriminante de q relativo a una base (a_i), es el determinante de las n formas lineales definidas por

$$(i = 1, \ldots, n) \qquad x \mapsto \frac{1}{2} q'_{x_i}.$$

II Formas degeneradas y no degeneradas. Ortogonalidad. Elementos isótropos

224. Formas bilineales simétricas degeneradas y no degeneradas

a) Núcleo de una forma bilineal simétrica

Sea f una *forma bilineal simétrica* sobre E, q la forma cuadrática asociada; designemos siempre por φ la aplicación lineal de E en E* definida por $\varphi(x) = f_x$. Se llama *núcleo* de f (o de q) el núcleo de φ, está constituido por los x, tales que

$$\varphi(x) = f_x = 0 \Leftrightarrow [(\forall\, y \in E) \qquad f_x(y) = f(x, y) = 0]$$

siendo f simétrico está igualmente descrito por y, tal que

$$\varphi(y) = f_y = 0 \Leftrightarrow [(\forall\, x \in E) \qquad f_y(x) = f(x, y) = 0].$$

OBSERVACION

Se observará que el núcleo de f *no es* la parte A de $E \times E$ descrita por (x, y) tal que $f(x, y) = 0$; se verificará a manera de ejercicio que A no es en general un subespacio vectorial de $E \times E$: esto se debe a que f es *bilineal y no lineal.*

DEFINICIÓN 2. — *Se dirá que la forma bilineal simétrica f (o la forma cuadrática asociada q) es* no degenerada *si y solamente si φ es inyectiva, es decir, si y solamente si* $\operatorname{Ker} f = \operatorname{Ker} \varphi = \{0\}$.

Si $\operatorname{Ker} f = \operatorname{Ker} \varphi \neq \{0\}$ *se dirá que f y q son degeneradas.*

Si E es de dimensión n, se llamará *rango* de f (o de q) el rango de φ, es decir, la dimensión de $\operatorname{Im} \varphi = \varphi(E)$.

En este caso (E de dimensión finita) φ es inyectiva si y solamente si φ es biyectiva, pues dim $E^* = \dim E = n$ (ver § 143, cor. 2, del t. 7). De donde

TEOREMA 6. — *Sea f una forma bilineal simétrica sobre* E, *espacio vectorial de dimensión n y* (α_{ij}) *la matriz asociada a f relativamente a una base de* E, *las propiedades siguientes son equivalentes:*

1. *f es no degenerada sobre* E.
2. $f(x, y) = 0$, *para todo x de* E, *implica* $y = 0$.
3. $f(x, y) = 0$, *para todo y de* E, *implica* $x = 0$.
4. *La aplicación* φ *de* E *en* E^*, *asociada a f, es un isomorfismo de espacios vectoriales.*
5. $\det (\alpha_{ij}) \neq 0$.

En particular si f es no degenerada sobre E, de dimensión finita, cualquiera que sea la forma lineal l^* de E^* existe x único de E, tal que $l^* = \varphi(x) = f_x$, es decir, *existe x único* tal que

$$(\forall y \in E) \qquad l^*(y) = f_x(y) = f(x, y).$$

b) **Identificación de E y de E* (E de dimensión finita) con la ayuda de una forma bilineal simétrica no degenerada**

Se ha *escogido* una forma bilineal simétrica f no degenerada sobre E, entre todos los isomorfismos de E sobre E^*, φ tiene un papel particular: podemos calificarle de *canónico relativamente a la forma f no degenerada escogida.* Entonces nos es posible gracias a φ identificar E y su dual E^* poniendo $f_x = \varphi(x) = x$ para todo x de E.

φ siendo *biyectiva* tendremos para todo x, todo y de E y todo y^* de E^* (ver § 222, *b*)

$$(1) \qquad \langle x, \varphi(y) \rangle = [\varphi(y)](x) = f_y(x) = f(x, y)$$

$$(1') \qquad \langle x, y^* \rangle = f(x, \varphi^{-1}(y^*))$$

y merced a la identificación $\varphi(y) = y$

$$(2) \qquad \langle x, y \rangle = f(x, y).$$

OBSERVACION

En la observación del § 150, *a*) y el ejercicio 171, fin del capítulo 7, hemos dicho que *no había en general* isomorfismos φ de E sobre E^* tal que para todo automorfismo u de E se tenga para todo x y todo y

$$\langle x, \varphi(y) \rangle = \langle u(x), (\varphi \circ u)(y) \rangle$$

Se ve que la igualdad anterior tiene lugar cuando φ es el isomorfismo asociado a la forma bilineal simétrica no degenerada f, *para todos los automorfismos u de* E *que verifican para todo x y todo y*

$$f(x, y) = f(u(x), u(y)).$$

Veremos en el párrafo 228 que estos automorfismos describen un grupo, el grupo ortogonal de f: no hay, por tanto, contradicción entre la observación del párrafo 150 y los resultados obtenidos aquí, φ no es canónico relativamente al grupo **GL**(E), *sino solamente* relativo al grupo ortogonal de f, que es un subgrupo de **GL**(E).

225. Elementos ortogonales, elementos isótropos relativos a una forma bilineal simétrica

E es un espacio vectorial sobre K, cuerpo conmutativo de característica distinta de 2, se designa por f una *forma bilineal simétrica* sobre E y por q la forma cuadrática asociada.

a) Ortogonalidad en E relativamente a f (o a q)

DEFINICIÓN 3. — *Se dice que x e y de* E *son ortogonales respecto a f (o a q) si* $f(x, y) = 0$.
Se dice que dos partes A *y* B *de* E, *son ortogonales respecto a f (o a q), si cada elemento de* A *es ortogonal a cada elemento de* B *respecto a f (o a q).*

OBSERVACION

Algunos autores emplean el término «elementos conjugados respecto a f (o a q)» en lugar de «elementos ortogonales relativamente a f (o a q)», nosotros no lo haremos debido a una razón que se aclarará en el subpárrafo *c*), más abajo.

Habiéndose escogido la forma f, diremos para simplificar "x e y *son ortogonales*" sobrentendiendo "respecto a f (o a q)". Podremos definir *dos subespacios* de E, sean F y G *ortogonales*: gracias a la bilinealidad de f, está claro que F y G son ortogonales si y solamente si una base de F y una base de G son ortogonales.

Siendo F un subespacio vectorial de E, el conjunto descrito por x, tal que

$$(\forall y \in F) \qquad f(x, y) = 0$$

es evidentemente un subespacio de E, en el que todos los elementos son ortogonales a F; se le llama *el subespacio vectorial* E *ortogonal a* F *respecto a f (o a q), y se le designa* $F^{\perp}$.

Si $F = Kx_0 (x_0 \neq 0)$, es decir, si F está engendrado por $x_0 \neq 0$, diremos que $F^{\perp}$ es el *subespacio ortogonal a* x_0.

Se observará la analogía de esta ortogonalidad en E respecto a f (o a q) con la ortogonalidad definida en el § 151, entre elementos de E y de E*: veremos la causa a continuación en *c*); notemos ahora que no hay que temer confusión alguna: en el § 151 F es un subespacio de E, $F^{\perp}$ es un subespacio de E*, y aquí $F^{\perp}$ es un subespacio de E.

Este hecho $F^{\perp} \subset E$, nos lleva a una noción que no existe en el estudio de la ortogonalidad entre elementos de E y de E* la de elementos isótropos.

b) Elementos y subespacios isótropos respecto a f (o a q)

DEFINICIÓN 4. — *Se dice que x, elemento de* E, *es* isótropo *respecto a f (o a q) si* $f(x, x) = q(x) = 0$.
Se dice que un subespacio, F *de* E, *es* isótropo *respecto a f (o a q) si existe $x \neq 0$ de* F, *ortogonal a todo* F.

Escogida la forma bilineal simétrica f, diremos "*x es isótropo*" o "F *es isótropo*" sobrentendido respecto a f (o a q).

Naturalmente el vector 0 es siempre isótropo, puede suceder que sea el único (ver ej. 1, a continuación). Mas generalmente si todo elemento del núcleo de f es isótropo, la recíproca no es cierta en general (ver ej. 2 último y § 230, cor. 2 del t. 14).

EJEMPLOS Y EJERCICIOS

1. Sea $E = \mathbf{R}^n$ referido a su base canónica, y provisto de la forma cuadrática q definida por

$$q(x) = (x^1)^2 + \dots + (x^n)^2$$

$q(x) = 0$ implica $x^1 = \dots = x^n = 0$, luego $x = 0$.

2. Sea $E = \mathbf{R}^4$ referido a su base canónica y provisto de la forma cuadrática q definida por

$$q(x) = (x^1)^2 + (x^2)^2 + (x^3)^2 - (x^4)^2$$

$a = (1, 0, 0, 1) \neq 0$ es isótropo. Dar las coordenadas de todos los vectores isótropos en función de tres parámetros reales. Buscar el núcleo de q y verificar que a no pertenece a este núcleo.

3. Sea $E = \mathbf{C}^n$ referido a su base canónica y provisto de la forma cuadrática q definida por

$$q(x) = (x^1)^2 + \dots + (x^n)^2$$

Para $n = 2$ $a = (1, i) \neq 0$ es isótropo.

Para $n = 3$ dar las coordenadas de todos los vectores isótropos en función de dos parámetros **complejos.**

4. Sea E un espacio vectorial sobre K, x_0 un elemento no nulo de E, busquemos en qué condición $F = Kx_0$ subespacio engendrado por x_0, es isótropo. Todos los elementos de F son de la forma $\lambda x_0 (\lambda \in K)$, luego F será isótropo si existe $\mu \neq 0$, tal que, cualquiera que sea λ, $f(\lambda x_0, \mu x_0) = \lambda\mu f(x_0, x_0) = 0$, esto deberá verificarse, en particular, para $\lambda = 1$, luego x_0 es isótropo, de donde:

El subespacio $Kx_0 (x_0 \neq 0)$ *es isótropo si y solamente si* x_0 *es isótropo.* Se dice entonces que Kx_0 es una *recta isótropa* (que pasa por 0, ver § 137).

5. Demostrar que toda familia de vectores no isótropos ortogonales dos a dos, es libre.

La noción de subespacio isótropo nos llevará a resultados interesantes relativos a F y a $F^\perp$. Sea F un subespacio de E, espacio vectorial de dimensión finita provista de una forma bilineal simétrica f, podemos considerar la restricción f_F de f a $F \times F$ (y la restricción q_F de q a F), esta restricción f_F es una forma bilineal simétrica sobre F, para todo x de F $(f_F)_x$ es un elemento del dual F^* de F, designemos por φ' la aplicación lineal de F en F^* definida por

$$\varphi'(x) = (f_F)_x.$$

Busquemos el núcleo de φ': está constituido por los elementos x de F, verificando

$$(\forall y \in F) \qquad [\varphi'(x)](y) = (f_F)_x(y) = f_F(x, y) = f(x, y) = 0$$

luego según la definición de $F^\perp$ dada más arriba

$$\operatorname{Ker} \varphi' = F \cap F^\perp.$$

Por lo tanto, $F \cap F^{\perp} = \{0\}$ es equivalente a φ' *inyectivo* y, por tanto, a f_F *no degenerado*: se dice que f *es no degenerada sobre* F. Esto significa también que el único vector de F ortogonal a todo F es 0; por lo tanto, que F *no es isótropo*.

Supongamos una de estas condiciones equivalentes realizada para todo x de E, la forma lineal f_x definida sobre E, puede ser considerada como una forma lineal definida sobre F, la designaremos también f_x; como f_F no es degenerada sobre F existe x_1 de F (ver § 224, *a*), consecuencia de la cuarta parte del teorema 6), tal que

$$(\forall y \in F) \qquad f_x(y) = (f_F)_{x_1}(y)$$

es decir, para todo y de F

$$f(x, y) = f(x_1, y) \Leftrightarrow f(x - x_1, y) = 0$$

dicho de otro modo $x_2 = x - x_1$ es ortogonal a F, luego para todo x de E existe x_1 de F y x_2 de $F^{\perp}$, tales que

$$x = x_1 + x_2$$

luego $E = F + F^{\perp}$, la hipótesis hecha: $F \cap F^{\perp} = \{0\}$ implica

$$E = F \oplus F^{\perp}$$

esta última relación implicando $F \cap F^{\perp} = \{0\}$ podemos enunciar:

TEOREMA 7. — *Siendo* E *un espacio vectorial de dimensión finita, para toda forma* f *bilineal simétrica sobre* E *y todo subespacio vectorial* F *de* E, *las propiedades son equivalentes:*

1. *La restricción de* f *a* F *es no degenerada.*
2. $F \cap F^{\perp} = \{0\}$.
3. F *no es isótropo.*
4. $E = F \oplus F^{\perp}$.

Se deduce inmediatamente:

COROLARIO 1. — *Si* F *es un subespacio vectorial no isótropo de* E, *de dimensión finita, se tiene*

$$\dim F^{\perp} = \dim E - \dim F.$$

COROLARIO 2. — *Siendo* E *de dimensión finita, y* x_0 *un vector isótropo de* E, *el subespacio ortogonal a* x_0 *es un hiperplano de* E *que no contiene a* x_0.

EJERCICIO

6. **Siendo** (a_i) $(1 \leq i \leq n)$ **una base de E, tal que,** $f(a_1, a_1) \neq 0$, **hallar la ecuación cartesiana del subespacio ortogonal** $F^{\perp}$ **a** $F = Ka_1$, **deducir que** $E = F \oplus F^{\perp}$ **sin utilizar el teorema 7.**

c) Caso en el que f es no degenerada sobre E de dimensión finita. Relación entre la ortogonalidad en E respecto a f y la ortogonalidad entre elementos de E y elementos de E*

La identificación de E y de E* gracias a la biyección φ (ver § 224, *b*) nos permite identificar las dos nociones de ortogonalidad:

1. Ortogonalidad entre x de E e $y = \varphi(y)$ de E*

$$\langle x, y\rangle = \langle x, \varphi(y)\rangle = \langle x, f_y\rangle = f_y(x) = f(x, y) = 0$$

2. Ortogonalidad entre x e y de E relativamente a f

$$f(x, y) = 0.$$

El ortogonal de F en E* y el ortogonal de F (en E) relativamente a f están confundidos y se designan por el mismo símbolo $F^{\perp}$, esto si f es no degenerada sobre E. Si f es degenerada sobre E no se puede hacer la identificación, hemos dicho que no hay ningún inconveniente en designar las dos ortogonales de F (en el sentido 1 y en el sentido 2) por el mismo símbolo: siendo uno subespacio de E*, el otro de E. Finalmente el isomorfismo φ nos permite extender a la ortogonalidad en el sentido 2 los resultados del § 151 relativos a la ortogonalidad en el sentido 1:

TEOREMA 8. — *Siendo f una forma bilineal simétrica no degenerada sobre* E, *de dimensión finita, y* $F^{\perp}$ *el ortogonal de* F *relativamente a f, para todo subespacio* F *de* E, *se tiene*

$$\dim F^{\perp} = \dim E - \dim F, \qquad (F^{\perp})^{\perp} = F.$$

Se tiene, pues,

$$F \cap F^{\perp} = \{0\} \Leftrightarrow (F^{\perp})^{\perp} \cap F^{\perp} = F \cap F^{\perp} = \{0\}$$

es decir, utilizando el teorema 7 .

COROLARIO 1. — *Si f es no degenerada sobre* E, $F^{\perp}$ *es no isótropo si y sólo si* F *es no isótropo.*

Finalmente el isomorfismo φ nos permitirá definir bases duales en E. A toda base $(a_i)(1 \leq i \leq n)$ de E se puede asociar una base única (a^{*i}) de E*, tal que para todo par (i, j) (ver § 150)

$$\langle a_i, a^{*j}\rangle = \delta_i^j$$

existe, pues, una base *única* (a_i^*) de E asociada a la base (a_i) por las fórmulas

$$(i = 1, \ldots, n) \qquad \varphi(a_i^*) = a^{*i}.$$

Para todo par (i, j), se tiene

$$f(a_i, a_j^*) = \langle a_i, \varphi(a_j^*)\rangle = \langle a_i, a^{*j}\rangle = \delta_{ij}.$$

Siendo simétrica la relación entre (a_i^*) y (a^{*i}), también lo es la relación entre (a^i) y (a_i^*) lo que justifica la definición siguiente:

DEFINICIÓN 5.—*Siendo f una forma bilineal simétrica no degenerada sobre* E, *espacio vectorial de dimensión finita, la base única* (a_i^*) *de* E, *asociada a la base* (a_i) *de* E, *por las fórmulas*

$$(i, j = 1, \ldots, n) \qquad f(a_i, a_j^*) = \delta_{ij}$$

se la llama base dual *de la base* (a_i).

Se dice también que las bases (a_i) *y* (a_i^*) *son* duales una de otra.

OBSERVACION

El lector debe percatarse de la diferencia entre los dos teoremas 7 y 8.

Por un lado f puede perfectamente ser no degenerada sobre E y degenerada sobre un subespacio vectorial F de E (ver ej. 7 más abajo).

Por otra parte, si f es no degenerada sobre E, la ortogonalidad en el sentido 1 no da ninguna información sobre $F \cap F^\perp$ (en el sentido 2): así la condición F no isótropo es siempre necesaria y suficiente para que $E = F \oplus F^\perp$, f sea degenerada o no.

EJERCICIOS

7. *a*) $E = \mathbf{R}^2$ referida a su base canónica, $f(x, y) = x^1y^1 - x^2y^2$, *estudiar la restricción de* f a $F = \mathbf{R}x_0$, $x_0 = (1, 1)$.

b) La misma pregunta con $E = \mathbf{C}^2$, $f(x, y) = x^1y^1 + x^2y^2$, $F = \mathbf{C}x_0$, $x_0 = (1, i)$.

8. Demostrar que en todos los casos $F \subset (F^\perp)^\perp$.

9. Determinar $F^\perp$, $(F^\perp)^\perp$ en los casos siguientes (las bases son siempre las bases canónicas).

$$E = \mathbf{R}^2,\ f(x, y) = x^1y^1 + x^2y^2, \qquad F = \mathbf{R}x_0,\ x_0 = (1, 0)$$
$$E = \mathbf{R}^2,\ f(x, y) = x^1y^1, \qquad F = \mathbf{R}x_0,\ x_0 = (1, 0)$$
$$E = \mathbf{R}^2,\ f(x, y) = x^1y^1, \qquad F = \mathbf{R}x_0,\ x_0 = (0, 1)$$
$$E = \mathbf{C}^2,\ f(x, y) = x^1y^1 + x^2y^2, \qquad F = \mathbf{C}x_0,\ x_0 = (1, i).$$

Comprobar que los resultados obtenidos están conformes con los teoremas 7 y 8.

226. Bases ortogonales y ortonormales. Formas canónicas

Siendo E un espacio vectorial de dimensión n sobre K de característica $\neq 2$, f es una forma bilineal simétrica sobre E, q la forma cuadrática asociada.

a) DEFINICIÓN 6.—*Una base* (a_i) $(1 \leq i \leq n)$ *de* E *es* ortogonal relativamente *a* f *(o a q) si*

$$i \neq j \Rightarrow f(a_i, a_j) = 0.$$

Una base (a_i) *de* E *es* ortonormal relativamente *a* f *(o a q) si es ortogonal y si*

$$(i = 1, \ldots, n) \qquad f(a_i, a_i) = 1.$$

Algunos autores dicen "ortonormada" por "ortonormal". Una base ortonormal es, pues, tal que para todo par (i, j) $f(a_i, a_j) = \delta_{ij}$; resulta de lo anterior que *una base ortonormal es dual de ella misma* (§ 225, def. 5).

EJEMPLO

La base canónica de K^n es ortonormal relativamente a f definida por

$$f(x, y) = \sum_{i=1}^{n} x^i y^i.$$

b) **Existencia de bases ortogonales. Forma canónica de $f(x, y)$ y $q(x)$ en una base ortogonal**

Si para x, $q(x) = f(x, x) = 0$, $f = 0$ (V. § 223, cor. 1 del t. 5) y todo par de vectores de E es un par de vectores ortogonales, luego toda base de E es ortogonal relativamente a $f = 0$.

Supongamos pues $f \neq 0$. Para $n = 1$, el problema de la existencia de una base ortogonal no se plantea. Si $n = 2$, existe a_1 tal que $f(a_1, a_1) = q(a_1) \neq 0$. Luego $F = Ka_1$ no es isótropo (V. § 225, ex. 4), por tanto $E = F \oplus F^\perp$ y F es de dimensión 1 (§ 225, t. 7) todo vector $a_2 \neq 0$ de $F^\perp$ forma con a_1 una base ortogonal de E. Si $n > 2$, supongamos (hipótesis de recurrencia) que existe una base ortogonal relativamente a f para todo espacio de dimensión $n - 1$. Como $f \neq 0$, existe a_1 tal que $f(a_1, a_1) = q(a_1) \neq 0$ por tanto $F = Ka_1$ no es isotropo, luego $E = F \oplus F^\perp$ (igual que más arriba) y $F^\perp$ es de dimensión $n - 1$, según la hipótesis de recurrencia existe una base $\{a_2, \ldots, a_n\}$ de $F^\perp$ ortogonal. Ahora bien a_i $(2 \leq i \leq n)$ es ortogonal a a_1 luego (a_i) $(1 \leq i \leq n)$ es una base ortogonal de E.

Relativamente a una base tal tendremos

$$f(x, y) = \sum_{i=1}^{n} \alpha_{ii} x^i y^i, \qquad q(x) = \sum_{i=1}^{n} \alpha_{ii} (x^i)^2$$

pues $i \neq j$ implica $\alpha_{ij} = f(a_i, a_j) = 0$.

Si todos los coeficientes α_{ii} son nulos $f = 0$; si s $(1 \leq s \leq n)$ los coeficientes son no nulos cambiando si es necesario la numeración de los vectores de la base ortogonal tendremos

$$f(x, y) = \sum_{i=1}^{s} \alpha_{ii} x^i y^i \qquad (i = 1, \ldots, s \qquad \alpha_{ii} \neq 0).$$

El núcleo de f estará descrito por $x = x^1 a_1 + \ldots + x^n a_n$ tal que

$$(\forall y \in E) \qquad f(x, y) = 0$$

es decir, por x tal que

$$(\alpha_{11} x^1 = \ldots = \alpha_{ss} x^s = 0) \Rightarrow (x^1 = \ldots = x^s = 0).$$

El núcleo de f está, pues, descrito por x verificando

$$x = x^{s+1} a_{s+1} + \ldots + x^n a_n$$

donde $x^{s+1}, \ldots, x^n$ son $n - s$ escalares arbitrarios, pues dim (Ker f) $= n - s$ y, en consecuencia, si r es el rango de f, $s = r$. Podemos, pues, enunciar:

TEOREMA 9. — *Para todo espacio vectorial* E *de dimensión finita n, existen bases ortogonales relativamente a toda forma bilineal simétrica f sobre* E.

De un modo más preciso, si f es de rango $r \neq 0$ existen unas bases, tales que

$$i \neq j \Rightarrow f(a_i, a_j) = 0$$
$$1 \leq i \leq r \Rightarrow f(a_i, a_i) \neq 0$$
$$r + 1 \leq i \leq n \Rightarrow f(a_i, a_i) = 0.$$

Con relación a una base, se tienen las formas canónicas

$$f(x, y) = \sum_{i=1}^{r} \alpha_{ii} x^i y^i, \qquad q(x) = \sum_{i=1}^{r} \alpha_{ii}(x^i)^2.$$

EJERCICIO

Demostrar el teorema 9 sin utilizar el teorema 7 (§ 225), pero utilizando el ejercicio 6 del párrafo 225.

OBSERVACIONES

1. Si relativamente a una base cualquiera (a_i) a f está asociada una matriz cuadrada A, relativamente a una *base ortogonal* (a'_i), la matriz A′ asociada a f es diagonal. Pero no se trata de la diagonalización definida en el capítulo 11, en efecto, siendo P la matriz de paso de (a_i) a (a'_i) se tiene $A' = {}^tPAP$.

Por otra parte si se considera A como asociada a la aplicación lineal φ, se tiene

$$A = M(\varphi, (a_i), (a^{*i})), \qquad A' = M(\varphi, (a'_i), (a'^{*i}))$$

luego A ha sido diagonalizada mediante un cambio de base en el espacio inicial E, de (a_i) a (a'_i), y de un cambio de base correlativo en el espacio de llegada E*; de (a^{*i}) a (a'^{*i}).

2. Si en una base cualquiera (a_i) se tiene (ver § 223)

$$q(x) = \sum_{i=1}^{n} \alpha_{ii}(x^i)^2 + 2 \sum_{1 \leq i < j \leq n} \alpha_{ij} x^i x^j$$

en una base ortogonal se tendrá

$$q(x) = \sum_{i=1}^{n} \alpha'_{ii}(x'^i)^2 = \sum_{i=1}^{n} \alpha'_{ii}(l^i(x))^2$$

poniendo

$$(i = 1, \ldots, n) \qquad x'^i = l^i(x) = \sum_{i=1}^{n} p^i_j x^j$$

las n formas l^i son independientes. Se tendrá, por tanto, si f es de rango r y si se supone $\alpha'_{11} \neq 0, \ldots, \alpha'_{rr} \neq 0$ y $\alpha_{r+1\, r+1} = \ldots = \alpha'_{nn} = 0$

$$q(x) = \sum_{i=1}^{r} \alpha'_{ii}(l^i(x))^2$$

se dice por abuso de lenguaje, que *q está descompuesto en cuadrados de r formas linealmente independientes.* El resultado del teorema 9, para que pueda ser un método de cálculo efectivo de la forma canónica necesita un cambio de base: daremos en el ejercicio 504 (fin del capítulo) un método directo de descomposición de una forma cuadrática en cuadrados linealmente independientes dada por GAUSS.

c) Existencia de bases ortonormales cuando f es no degenerada y cuando $K = C$. Formas canónicas correspondientes

Si existe una base ortonormal relativamente a f, se tiene para todo par (i, j), $f(a_i, a_j) = \delta_{ij}$ luego para todo i, $\alpha_{ii} = 1$ y por consecuencia $r = n$ y f es no degenerada.

Sea, pues, f una *forma no degenerada* sobre E, E es un espacio vectorial de dimensión n sobre **C**. Existen las bases ortogonales (b_i) y se tiene

$$i \neq j \Rightarrow f(b_i, b_j) = 0$$
$$1 \leq i \leq n \Rightarrow f(b_i, b_i) = \beta_{ii} \neq 0.$$

Al ser todo elemento un cuadrado en **C** pongamos

$$(i = 1, \ldots, n) \qquad (\lambda_i)^2 = \beta_{ii}, \qquad a_i = \frac{b_i}{\lambda_i}$$

tendremos

$$i \neq j \Rightarrow f(a_i, a_j) = 0$$
$$1 \leq i \leq n \Rightarrow f(a_i, a_i) = \frac{f(b_i, b_i)}{(\lambda_i)^2} = 1$$

donde $\beta_{ii} \neq 0$ implica $\lambda_i \neq 0$, de donde:

TEOREMA 10. — *Si* E *es un espacio vectorial de dimensión n sobre* **C** *y f una forma bilineal simétrica no degenerada sobre* E, *existen bases de* E *ortonormales relativamente a f. En relación a una base tal se tienen las formas canónicas*

$$f(x, y) = \sum_{i=1}^{n} x^i y^i, \qquad q(x) = \sum_{i=1}^{n} (x^i)^2.$$

OBSERVACIONES

1. Siendo f una forma bilineal simétrica no degenerada sobre E, espacio vectorial sobre K, existirán bases ortonormales si en K todo elemento es un cuadrado; por ejemplo, si K es algebraicamente cerrado. Existirá igualmente bases ortonormales relativamente a f, si existen bases ortogonales (a_i), tales que para $i = 1, \ldots, n$, $f(a_i, a_i)$ sea un cuadrado en K (ver § 230).

2. Siendo E de dimensión n sobre **C** hay, pues, *una sola forma canónica* para toda forma bilineal simétrica no degenerada sobre E.

III. Endomorfismo adjunto. Aplicaciones

227. Endomorfismo adjunto

a) Introducción

En toda esta sección, f será una forma bilineal simétrica no degenerada sobre E, *espacio vectorial de dimensión finita sobre* K, *cuerpo conmutativo de* característica $\neq 2$; q designará la forma cuadrática asociada a f y φ el isomorfismo de E sobre E* asociado a f.

Designaremos por u un endomorfismo de E, es decir, un elemento de $\mathcal{L}(E)$ y por ${}^t u$ la transpuesta de la aplicación lineal u, es decir, el elemento único de $\mathcal{L}(E^*)$ definido por (ver § 152 tomando $F = E$)

$$(\forall\, x \in E) \qquad (\forall\, y^* \in E^*) \qquad \langle u(x),\ y^* \rangle = \langle x,\ {}^t u(y^*) \rangle.$$

Por otra parte, si u y v son endomorfismos de E, consideraremos a menudo las relaciones de la forma

$$(\forall (x,\ y) \in E^2) \qquad f(u(x),\ y) = f(v(x),\ y)$$

Por tanto, para todo y

$$f_{u(x)}(y) = f_{v(x)}(y)$$

es decir, $f_{u(x)} = f_{v(x)}$ para todo x; ahora bien, $f_{u(x)} = \varphi[u(x)]$ y $f_{v(x)} = \varphi[v(x)]$, entonces siendo φ biyectiva

$$u(x) = v(x)$$

y esto para todo x; en consecuencia, la relación dada implica $u = v$.

b) Adjunto de un endomorfismo relativamente a f (o bien a q)

Si φ es el isomorfismo de E y E* definido por $\varphi(x) = f_x$ tendremos (ver § 224, *b*, fórmulas (1) y (1′))

$$f(u(x),\ y) = \langle u(x),\ \varphi(y) \rangle = \langle x,\ {}^t u[\varphi(y)] \rangle$$
$$= \langle x,\ ({}^t u \circ \varphi)(y) \rangle = f(x,\ (\varphi^{-1} \circ {}^t u \circ \varphi)\ (y))$$

y esto cualesquiera que sean x e y. Al endomorfismo u, f permite, por tanto, asociar el endomorfismo único $\varphi^{-1} \circ {}^t u \circ \varphi$; el diagrama

$$E \xrightarrow{\varphi} E^* \xrightarrow{{}^t u} E^* \xrightarrow{\varphi^{-1}} E$$

demuestra que $\varphi^{-1} \circ {}^t u \circ \varphi$ es un endomorfismo de E, se le llama el *adjunto* de u relativamente a f y se le designa u^*. (Atención: A pesar de esta notación, es un elemento de $\mathcal{L}(E)$ y no de E*.) El adjunto de u está, por tanto, definido por

$$(1) \qquad (\forall (x,\ y) \in E^2) \qquad f(u(x),\ y) = f(x,\ u^*(y)).$$

Si se identifica E y E* gracias al isomorfismo φ poniendo $\varphi(x) = x$(§ 224, *b*) *las propiedades de* ${}^t u$ *se extienden inmediatamente a* u^*. Se tiene, pues (ver § 152), cualquiera que sean u y v de $\mathcal{L}(E)$ y λ de K

$$(u + v)^* = u^* + v^*, \qquad (\lambda u)^* = \lambda u^*, \qquad (u \circ v)^* = v^* \circ u^*$$
$$\text{rg } (u^*) = \text{rg } (u), \qquad (u^*)^* = u$$

si, además, u es inversible, u^* lo es también

$$(u^*)^{-1} = (u^{-1})^*$$

naturalmente

$$(\text{id}_E)^* = \text{id}_E.$$

Todos estos resultados podrían, por otra parte, demostrarse directamente mediante la fórmula (1).

Finalmente el diagrama utilizado anteriormente, completado por la indicación de las bases empleadas

$$u^* \ : \ \underset{(a_i^*)}{E} \xrightarrow{\varphi} \underset{(a^{*i})}{E^*} \xrightarrow{{}^t u} \underset{(a^{*i})}{E^*} \xrightarrow{\varphi^{-1}} \underset{(a_i^*)}{E}$$

siendo (a_i^*) la base de E dual de (a_i) y (a^{*i}) la base de E* dual de la base (a_i) de E, demuestra que (ver § 155, *b*) y fórmula final del § 225: $\varphi(a_i^*) = (a^{*i})$

$$M(u^*, (a_i^*)) = I_n M({}^t u, (a^{*i})) I_n = {}^t M(u, (a_i))$$

por tanto (ver § 169, *c*)

$$\det (u^*) = \det (u).$$

En particular, si existe para E una base ortonormal, relativamente a f, es dual de sí misma; en consecuencia,

$$M(u^*, (a_i)) = {}^t[M(u, (a_i))].$$

Resumamos todos estos resultados:

TEOREMA 11. — *Sea* E *un espacio vectorial de dimensión finita, f una forma bilineal simétrica no degenerada sobre* E *y u un endomorfismo cualquiera de* E, *existe un endomorfismo único u* de* E, *llamado* adjunto de u relativamente *a f, tal que*

(1) $$(\forall (x, y) \in E^2) \qquad f(u(x), y) = f(x, u^*(y)).$$

Cualesquiera que sean u y v de $\mathcal{L}$(E) *y* λ *de* K

$$(u + v)^* = u^* + v^*, \qquad (\lambda u)^* = \lambda u^*, \qquad (u \circ v) = v^* \circ u^*$$
$$(u^*)^* = u, \qquad \text{rg } (u^*) = \text{rg } (u), \qquad \det (u^*) = \det (u).$$

Si u es inversible, u lo es también y* $(u^*)^{-1} = (u^{-1})^*$. *Las matrices de u relativamente a una base de* E *y de u* respecto a la base dual, relativamente a f, son transpuestas una de otra; en particular las matrices de u y de u* respecto a una base ortonormal relativamente a f, si existe, son transpuestas una de otra.*

EJERCICIO

Si existe una base ortonormal (a_i), demostrar que en esta base $M(u^*) = {}^t(M(u))$ utilizando la relación (1).

228. Grupo ortogonal GL(*f*) o GL(*q*)

a) Operadores ortogonales. Matrices ortogonales

Supongamos que u conserve f (o q lo que es equivalente, § 223, cor. 2 del t. 5), es decir,

$$(2) \qquad (\forall\ (x,\ y) \in E^2) \qquad f(u(x),\ u(y)) = f(x,\ y)$$

vamos a ver que u es un automorfismo de E y que $u^{-1} = u^*$. Tenemos, en efecto, utilizando (1)

$$f(x,\ y) = f(u(x),\ u(y)) = f(x,\ u^*[u(y)])$$

es decir,

$$(3) \qquad (\forall (x,\ y) \in E^2) \qquad f(x,\ y) = f(x,\ (u^* \circ u)\ (y)).$$

Utilizando la equivalencia demostrada al principio del § 227 vemos que (3) implica

$$(4) \qquad u^* \circ u = \mathrm{id}_E.$$

Recíprocamente (4) implica (3); por tanto, (2). La relación (4) implica que u es inyectiva (ver nota 38, pie de página) y, en consecuencia, biyectiva, puesto que E es de dimensión finita (§ 143, cor. 2 del t. 7), luego $u^* = u^{-1}$ y se tiene también

$$(4') \qquad u \circ u^* = \mathrm{id}_E.$$

Supongamos que se verifica la relación (4′), u es entonces *suprayectiva*(38), por tanto, biyectiva (§ 143, cor. 2 del t. 7), siendo E de dimensión finita. Resumamos los resultados obtenidos:

TEOREMA 12. — *Sea* E *un espacio vectorial de dimensión finita, y f una forma bilineal simétrica no degenerada sobre* E, *para todo endomorfismo u de* E, *las cinco propiedades siguientes son equivalentes:*

1.ª $(\forall\ (x,\ y) \in E^2) \qquad f[u(x),\ u(y)] = f(x,\ y)$.
2.ª $(\forall\ x \in E) \qquad q[u(x)] = q(x)$.
3.ª $u^* \circ u = \mathrm{id}_E$.
4.ª $u \circ u^* = \mathrm{id}_E$.
5.ª *u es inversible y* $u^{-1} = u^*$.

Un endomorfismo verificando una de estas cinco propiedades recibe el nombre de *automorfismo ortogonal* de E, relativamente a f (o a q). Se le dice

(38) Si f y g son dos aplicaciones de un conjunto A en sí mismo, verificando $g \circ f = \mathrm{id}_A$, f es *inyectiva*, pues $f(x) = f(x')$ implica $x = g[f(x)] = g[f(x')] = x'$; g es suprayectiva, pues si no fuera suprayectiva la relación $g[f(x)] = x$ no se verificaría para x perteneciendo a $A - g(A) \neq \varnothing$. Estas propiedades son casos particulares de las propiedades enunciadas en el ejercicio 1 del § 15.

también endomorfismo u *operador ortogonal* (pues el hecho de ser ortogonal implica el que sea un automorfismo).

Se dice también que u es un *automorfismo* de f (o de q).

La propiedad (1) implica que si la base (a_i) es ortonormal, y si u es ortogonal, la familia $(u(a_i))$ es una base ortonormal; en efecto, para todo par (i, j)

$$f[u(a_i), u(a_j)] = f(a_i, a_j) = \delta_{ij}$$

recíprocamente si esta relación es satisfecha u es ortogonal, ya que, en efecto, poniendo

$$x = \sum_{i=1}^{n} x^i a_i$$

$$f[u(x), u(x)] = f\left[\sum_{i=1}^{n} x^i u(a_i), \sum_{j=1}^{n} x^j u(a_j)\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} x^i x^j f[u(a_i), u(a_j)]$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} x^i x^j f(a_i, a_j) = \sum_{i=1}^{n} (x^i)^2 = q(x)$$

y según la propiedad (2) u es ortogonal, de donde: (a_i) *es una base ortonormal de* E, *relativamente a* f, *el operador* u *es ortogonal si y sólo si la familia* $(u(a_i))$ *es una base ortonormal de* E *relativamente a* f.

Sea A la matriz asociada a un operador ortogonal respecto a una base ortonormal relativamente a f, si las hay se tendrá

$${}^t\mathrm{A}\mathrm{A} = \mathrm{A}\,{}^t\mathrm{A} = \mathrm{I}_n \Leftrightarrow \mathrm{A}^{-1} = {}^t\mathrm{A}.$$

De una manera general una matriz de $\mathbf{M}_n(\mathrm{K})$ verificando una de estas condiciones se la llama matriz *ortogonal*. Si se pone $\mathrm{A} = (\alpha_i^j)$ se tendrá

$$\begin{cases} i \neq j \Rightarrow \quad \sum_{k=1}^{n} \alpha_j^k \alpha_i^k = 0 \\ i = 1, \dots, n \quad \sum_{k=1}^{n} (\alpha_i^k)^2 = 1 \end{cases}$$

escribiendo ${}^t\mathrm{A}\mathrm{A} = \mathrm{I}_n$; se obtendría un sistema equivalente escribiendo $\mathrm{A}\,{}^t\mathrm{A} = \mathrm{I}_n$.

En particular si existe en E bases ortonormales, relativamente a f, *toda matriz de paso de una base ortonormal a una base ortonormal es ortogonal.*

En efecto, pongamos (ver § 160).

$$\mathrm{X} = \mathrm{M}(x, (a_i)), \qquad \mathrm{X}' = \mathrm{M}(x, (a'_i)), \qquad \mathrm{X} = \mathrm{P}\mathrm{X}'$$

tenemos con (a_i) y (a'_i) ortonormales

$$f(x, x) = q(x) = \sum_{i=1}^{n} (x^i)^2 = {}^t\mathrm{X}\mathrm{X} = \sum_{i=1}^{n} (x'^i)^2 = {}^t\mathrm{X}'\mathrm{X}'$$

de donde
$$q(x) = {}^t(\mathrm{PX'})(\mathrm{PX'}) = {}^t\mathrm{X'}({}^t\mathrm{PP})\mathrm{X'} = {}^t\mathrm{X'X'}$$
de donde (ver § 221, t. 2)
$${}^t\mathrm{PP} = \mathrm{I}_n.$$

b) Grupo ortogonal (relativamente a *f*)

Es fácil ver que los automorfismos ortogonales de E relativamente a f (o a q) describen un subgrupo de $\mathbf{GL}_n(\mathrm{K})$; en efecto, (ver § 72, *c*),
$$(u^{-1} = u^* \quad \text{y} \quad v^{-1} = v^*) \Leftrightarrow (u \circ v^{-1})^{-1} = (u \circ v^{-1})^*$$
pues
$$(u \circ v^{-1})^{-1} = v \circ u^{-1} = (v^{-1})^* \circ u^* = (u \circ v^{-1})^*$$
de donde:

COROLARIO. — *El conjunto de los automorfismos ortogonales relativamente a f(o a q) es un subgrupo de* $\mathbf{GL}_n(\mathrm{K})$, *llamado* grupo ortogonal *relativamente a f (o a q) y representado* $\mathbf{GL}(f)$ *o* $\mathbf{GL}(q)$.

Si existen bases ortonormales en E (relativamente a f) hemos visto que sólo existe una única forma canónica para $f(x, y)$ (o $q(x)$)
$$f(x, y) = \sum_{i=1}^{n} x^i y^i, \qquad q(x) = \sum_{i=1}^{n} (x^i)^2$$
el grupo ortogonal relativo a esta forma se designa $\mathbf{O}(n, \mathrm{K})$.

Naturalmente las matrices de orden n ortogonales, con elementos en K, describen un subgrupo del grupo de las matrices inversibles de $\mathbf{M}_n(\mathrm{K})$ que es isomorfo a $\mathbf{O}(n, \mathrm{K})$.

Por otra parte, para todo operador ortogonal se tiene
$$\det\,(u^* \circ u) = (\det\, u)^2 = \det\,(\mathrm{id}_\mathrm{E}) = 1$$
1 designando el elemento unidad de K.

Si $u^{-1} = u^*$ y $\det\,(u) = +1$ se dice que u es una *rotación* de E (relativamente a f o q); está claro que las rotaciones de E describen un subgrupo de $\mathbf{GL}(f)$ llamado *grupo de rotaciones* de E (relativamente a f o a q) o bien también *grupo ortogonal especial*.

El conjunto de las matrices ortogonales A tales que $\det \mathrm{A} = +1$ describen un grupo designado $\mathbf{O}^+(n, \mathrm{K})$ o $\mathbf{SO}(n, \mathrm{K})$ naturalmente isomorfo al grupo de las rotaciones de E.

EJERCICIOS

1. Demostrar que el grupo de las rotaciones de E, relativamente a f, es un subgrupo distinto de $\mathbf{GL}(f)$. (Considerar el homomorfismo $u \to \det\,(u)$).
2. Poner de manifiesto las relaciones verificadas por los elementos de las matrices $\mathbf{O}(n, \mathrm{K})$ y de $\mathbf{O}^+(n, \mathrm{K})$ para $n = 2$ o 3.
3. Si λ de K, es un valor propio de un operador ortogonal, demostrar que $\lambda^2 = 1$.

229. Operadores simétricos relativamente a f (o a q)

a) Operadores simétricos

Busquemos los operadores *auto-adjuntos* relativamente a f, es decir, tales que $u = u^*$. Se tendrá

(5) $$(\forall (x, y) \in E^2) \qquad f[u(x), y] = f[x, u(y)]$$

recíprocamente (5) implica para todo par de E^2

$$f[x, u(y)] = f[u(x), y] = f[x, u^*(y)]$$

por tanto, utilizando la equivalencia demostrada al principio del § 227, $u = u^*$.

Respecto a una base ortonormal (a_i) relativamente a f, si las hay, se tendrá (ver § 227, t. 11)

$$u = u^* \Leftrightarrow M(u, (a_i)) = {}^t[M(u, (a_i))]$$

de donde:

TEOREMA 13. — *Sea* E *un espacio vectorial de dimensión finita, y f una forma bilineal simétrica no degenerada sobre* E, *para todo endomorfismo u de* E *las propiedades siguientes son equivalentes* (u^* *es el adjunto de u relativamente a f*):

1.ª $u = u^*$.

2.ª $(\forall (x, y) \in E^2) \qquad f[u(x), x] = f[x, u(y)]$.

3.ª *Si existen bases ortonormales, relativamente a f, la matriz asociada a u en una tal base es simétrica.*

Un endomorfismo de E verificando una de estas propiedades se llama *operador simétrico relativamente a f*.

b) Estudio de un isomorfismo de espacios vectoriales

Por otro lado, algunos de los resultados enunciados en el teorema 11 (§ 227) demuestran que el conjunto de los operadores simétricos relativamente a f es un *subespacio vectorial* de $\mathfrak{L}(E)$; lo designaremos $\mathfrak{S}_f(E)$.

A todo elemento u de $\mathfrak{S}_f(E)$ la fórmula

$$[\forall (x, y) \in E^2] \qquad f_u(x, y) = f(u(x), y)$$

permite asociar una aplicación de E^2 en K, se ve inmediatamente que es *bilineal* y *simétrica*; en consecuencia, f_u es un elemento del espacio vectorial $\mathfrak{S}_2(E, K)$ de las formas bilineales simétricas sobre E. Mediante procedimientos análogos a los utilizados en el § 222, *b*) se demuestra que $\mathfrak{S}_f(E)$ y $\mathfrak{S}_2(E, K)$ son isomorfos (ver ej. 520 final del capítulo).

Si existen bases ortonormales respecto a f este isomorfismo es evidente: sea (a_i) una base ortonormal, poniendo $X = M(x, (a_i))$, $Y = M(y, (a_i))$ $A = M[u, (a_i)]$, (${}^tA = A$) tendremos

$$f(x, y) = {}^tYX, \qquad f[u(x), y] = {}^tY(AX) = {}^tYAX$$

Por otra parte, sea g un elemento de $\mathfrak{S}_2(E, K)$ al que está asociada respecto a (a_i) la matriz $B({}^tB = B)$, tendremos

$$g(x, y) = {}^tYBX.$$

En consecuencia, *en la base* (a_i)

$$g = f_u \Leftrightarrow A = B$$

la aplicación $u \to f_u$, así definida, gracias a una base ortonormal (a_i) es independiente de esta base; en efecto, sea P la matriz de paso de (a_i) a otra base ortonormal (a'_i), P es una matriz ortogonal (ver § 228, *a*), por tanto $P^{-1} = {}^tP$, en la nueva base A′, que es asociada a *u*, y B a g

$$A' = P^{-1}AP, \qquad B' = {}^tPBP = A'$$

Designado por S(*n*, K) el espacio vectorial de las matrices cuadradas simétricas de orden *n* con elementos en K el isomorfismo de $\mathcal{S}_f(E)$ sobre $\mathcal{S}_2(E;\ K)$ es el compuesto de los isomorfismos

$$\mathcal{S}_f(E) \to S(n,\ K) \to \mathcal{S}_2(E;\ K).$$

IV. Formas bilineales simétricas reales. Espacio euclídeo de dimensión *n*

Supondremos en toda esta sección que E es un espacio vectorial sobre **R**; diremos que *f*, forma bilineal simétrica sobre E, y *q*, forma cuadrática asociada, son *reales*.

230. Primeras propiedades

a) Formas bilineales simétricas positivas. Desigualdad de Schwarz

Definición 7. — *Diremos que una forma f, bilineal simétrica real sobre* E, *y que la forma cuadrática asociada q son* positivas *si*

$$(\forall x \in E) \qquad f(x,\ x) = q(x) \geq 0.$$

Se definiría igualmente las formas negativas ($q(x) \leq 0$), donde *f* (o *q*) es negativa si y sólo si $-f$ (o $-q$) es positiva: estudiaremos sólo formas positivas.

Siendo *f* positiva, para todo par $(x,\ y)$ de E^2 y todo número real λ tendremos

$$q(x + \lambda y) = f(x + \lambda y,\ x + \lambda y) = f(x,\ x) + 2\lambda f(x,\ y) + \lambda^2 f(y,\ y) \geq 0$$

en el caso en que $f(y,\ y) \neq 0$, $q(x + \lambda)$ es un polinomio de segundo grado en λ con coeficientes reales; es siempre ≥ 0 y es, por tanto, imposible que tenga dos raíces reales distintas, de donde

$$(1) \qquad [f(x,\ y)]^2 - f(x,\ x)\,f(y,\ y) \leq 0.$$

Si *y* es tal que $f(y,\ y) = 0$, la desigualdad $q(x + \lambda y) \geq 0$ no se puede verificar para todo real λ más que si $f(x,\ y) = 0$, luego (1) se verifica también, de donde:

TEOREMA 14 *(desigualdad de Schwarz).* — *Si f es una forma bilineal simétrica positiva sobre* E, *se tiene*

(1) $$(\forall (x, y) \in E^2) \qquad [f(x, y)]^2 - f(x, x)f(y, y) \leq 0.$$

Se deduce de este teorema, para todo par (x, y) de E^2

$$q(x + y) = f(x + y, x + y) = f(x, x) + 2f(x, y) + f(y, y)$$
$$\leq q(x) + 2\sqrt{q(x)q(y)} + q(y) = [\sqrt{q(x)} + \sqrt{q(y)}]^2$$

pues (1) implica: $f(x, y) \leq \sqrt{q(x)q(y)}$, de donde:

COROLARIO 1. — *Si q es una forma cuadrática positiva sobre* E, *se tiene*

(2) $$(\forall (x, y) \in E^2) \qquad \sqrt{q(x + y)} \leq \sqrt{q(x)} + \sqrt{q(y)}.$$

Se sabe, por otra parte, que el núcleo de f (§ 224, *a*) está descrito por x tal que

$$(\forall y \in E) \qquad f(x, y) = 0$$

por tanto (ver § 225, *b*), todo elemento del núcleo es isótropo. Vamos a demostrar recíprocamente que, cuando *f es positiva,* todo elemento isótropo, respecto a f, pertenece al núcleo; en efecto, $f(x, x) = 0$ implica, gracias a la desigualdad de SCHWARZ, $[f(x, y)]^2 \leq 0$ y, en consecuencia, $f(x, y) = 0$ para todo y, de donde:

COROLARIO 2. — *El núcleo de una forma f bilineal simétrica positiva sobre* E *es el conjunto de los vectores isótropos de* E, *relativamente a f.*

De lo anterior resulta que f, positiva, es no degenerada el núcleo se reduce a $\{0\}$; por tanto, *$x = 0$ es el único vector isótropo relativamente a f, forma bilineal simétrica positiva no degenerada sobre* E.

OBSERVACION

Algunos autores llaman *formas cuadráticas «definidas positivas»* (resp. negativas), las formas cuadráticas no degeneradas positivas (resp. negativas), aquí no lo haremos, los términos *no degenerados positivos* (resp. negativas) son suficientes.

EJEMPLOS Y EJERCICIOS

1. Sea $E = \mathbf{R}^n$ la forma definida (respecto a la base canónica de $\mathbf{R}^n$) por la fórmula

$$f(x, y) = \sum_{i=1}^{n} x^i y^i$$

es no degenerada positiva.

La forma definida en las mismas condiciones por

$$g(x, y) = \sum_{i=1}^{m} x^i y^i \qquad (m < n)$$

es positiva, pero es degenerada.

2. La forma bilineal simétrica f, definida en el ejercicio 2 del § 222 es positiva, se tiene, pues, $(\alpha < \beta)$

$$\left[\int_\alpha^\beta x(t)y(t)dt\right]^2 \leq \int_\alpha^\beta [x(t)]^2dt \int_\alpha^\beta [y(t)]^2dt$$

$$\left[\int_\alpha^\beta [x(t) + y(t)]^2dt\right]^{1/2} \leq \left[\int_\alpha^\beta [x(t)]^2dt\right]^{1/2} + \left[\int_\alpha^\beta [y(t)]^2dt\right]^{1/2}$$

Demostrar que f es no degenerada.

b) Caso en que E es de dimensión finita. Ley de inercia

Sea f una forma bilineal simétrica real sobre E, existen en E bases (b_i) $(1 \leq i \leq n)$ ortogonales relativamente a f (§ 226, t. 9). $i \neq j$ implica $f(b_i, b_j) = 0$; $f(b_i, b_i)$ real, no nulo. Si f es no nulo, cambiando si fuera necesario la numeración de la base (b_i), existe, por tanto, enteros s y t no nulos los dos tales que

$$\begin{array}{ll} 1 \leq i \leq s & f(b_i, b_i) = \beta_{ii} > 0 \\ s+1 \leq j \leq s+t & f(b_j, b_j) = \beta_{jj} < 0 \\ s+t+1 \leq k \leq n & f(b_k, b_k) = \beta_{kk} = 0 \end{array}$$

naturalmente $s + t = r = \text{rg}\ (f)$; si $t = 0$ (resp. $s = 0$) la forma f es positiva (resp. negativa). Volvamos al caso general, pongamos con las notaciones anteriores

$$a_i = \frac{b_i}{\sqrt{\beta_{ii}}}, \qquad a_j = \frac{b_j}{\sqrt{-\beta_{jj}}}, \qquad a_k = b_k$$

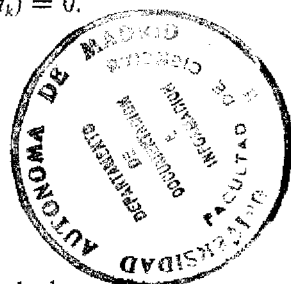
la base $(a_1, \ldots, a_n)$ es aún ortogonal, además

$$f(a_i, a_i) = 1, \qquad f(a_j, a_j) = -1, \qquad f(a_k, a_k) = 0.$$

Tenemos, por tanto, en la base (a_i) $(1 \leq i \leq n)$

$$f(x, y) = \sum_{i=1}^{s} x^i y^i - \sum_{j=s+1}^{s+t} x^j y^j$$

$$q(x) = \sum_{i=1}^{s} (x^i)^2 - \sum_{j=s+1}^{s+t} (x^j)^2.$$



Se dice algunas veces que $\beta_{ii}(x^i)^2$ $(\beta_{ii} > 0)$ y $(x^i)^2$ son «cuadrados positivos» y que $\beta_{jj}(x^j)^2$ $(\beta_{jj} < 0)$ y $-(x^j)^2$ son «cuadrados negativos»: en el primer caso esta denominación es una trivialidad, en el segundo caso es un abuso de lenguaje.

Supongamos que existe una segunda base (a'_i) ortogonal relativamente a f, y tal que respecto a ella se tenga

$$q(x) = \sum_{i=1}^{i=s'} (x'_i)^2 - \sum_{j=s'+1}^{j=s'+t'} (x'^j)^2.$$

Vamos a demostrar que $s' = s$ y $t = t'$. Consideremos los dos subespacios de E siguientes:

F engendrado por $a_1, \ldots, a_s$ (dim F $= s$)
G' engendrado por $a'_{s'+1}, \ldots, a'_n$ (dim G' $= n - s'$).

Si $x \in$ F se tiene $q(x) \geq 0$; si, además, $x \neq 0$ existe al menos un $x^i \neq 0$ $(1 \leq i \leq s)$, luego $q(x) > 0$ para todo $x \neq 0$ de F.

Si $x \in$ G', $q(x) \leq 0$, resulta que

$$F \cap G' = \{0\}$$

la fórmula demostrada en el ejercicio 2 del § 137 da

$$\dim F + \dim G' = \dim (F + G') \leq \dim E$$

de donde

$$s + n - s' \leq n \Leftrightarrow s \leq s'$$

permutando el papel de las bases (a_i) y (a'_i) se demostraría $s' \leq s$, de donde $s = s'$; como, por otro lado, $s + t = s' + t' = r = \text{rg}\ (f)$, se tiene también $t = t'$, de donde:

TEOREMA 15. — *Si q es la forma cuadrática asociada a una forma f bilineal simétrica real sobre* E *espacio vectorial de dimensión n sobre* **R**.

1. *Existe una base (a_i) ortogonal relativamente a f y dos enteros naturales s y t, tales que respecto a esta base (a_i), se tiene*

$$q(x) = \sum_{i=1}^{i=s} (x^i)^2 - \sum_{j=s+1}^{j=s+t} (x^j)^2 \qquad (s + t = r = rg(f).$$

2. *Si existe una segunda base (a'_i) y dos enteros naturales s' y t' tales que respecto a (a'_i), se tiene*

$$q(x) = \sum_{i=1}^{s'} (x'^i)^2 - \sum_{j=s'+1}^{s'+t'} (x'^j)^2 \qquad (s' + t' = r = rg(f))$$

entonces $s = s'$, $t = t'$ (ley de inercia).

Al par de enteros naturales (s, t) se le llama la *signatura* de f (o de q). Por ejemplo, la forma cuadrática sobre $\mathbf{R}^4$ definida por

$$q(x) = (x^1)^2 + (x^2)^2 + (x^3)^2 - (x^4)^2$$

tiene por signatura (3, 1).

Está claro que f forma bilineal simétrica real es positiva si y sólo si $t = 0$; por tanto, $s = r = \text{rg}\ (f)$, y que f es *no degenerada positiva* si y solamente si $s = n$, se tiene entonces respecto a la base (a_i)

$$f(x, y) = \sum_{i=1}^{n} x^i y^i, \qquad q(x) = \sum_{i=1}^{n} (x^i)^2.$$

COROLARIO 1. — *Si* E *es un espacio vectorial de dimensión finita sobre* **R**, *existen bases de* E *ortonormales relativas a toda forma bilineal simétrica no degenerada positiva sobre* E.

231. Espacio euclídeo de dimensión *n*

a) Noción de espacio vectorial normado

DEFINICIÓN 8. — *Sea* E *un espacio vectorial sobre* $K = \mathbf{R}$ o **C**, *toda aplicación* p *de* E *en* $\mathbf{R}_+$ *poseyendo las propiedades siguientes*

N_1. $p(x) = 0 \Rightarrow x = 0$

N_2. $(\forall (x, y) \in E^2) \quad p(x + y) \leq p(x) + p(y)$

N_3. $(\forall x \in E) (\forall \lambda \in K) \quad p(\lambda x) = |\lambda| p(x)$

se llama norma *sobre* E. *El espacio vectorial* E *provisto de una norma se le llama* espacio vectorial normado.

Se puede reemplazar **R** o **C** en esta definición por todo *cuerpo valorado* K (ver § 110, def. 2), $\lambda \to |\lambda|$ es entonces el valor absoluto definido sobre K.

En lugar de $p(x)$ se escribe $\|x\|$, o bien, $\|x\|_1$, $\|x\|_2$, ... si están definidas varias normas sobre E.

EJEMPLOS Y EJERCICIOS

1. **Q**, **R**, **C** o todo cuerpo valorado, considerado como espacio vectorial sobre sí mismo es un espacio vectorial normado (demostrar que $x \to |x|$ es una norma).

2. Demostrar que en $E = \mathbf{R}^n$ descrito por $x = (x^1, ..., x^n)$ las aplicaciones siguientes de E en $\mathbf{R}_+$ son normas

$$x \to \sum_{i=1}^{n} |x^i|, \qquad x \to \sup_{1 \leq i \leq n} |x^i|.$$

3. Siendo E un espacio vectorial normado, demostrar que la aplicación de $E \times E$ en $\mathbf{R}_+$ definida por $(x, y) \to \|x - y\|$ es una distancia sobre E (ver § 10, def. 3).

4. Siendo E el espacio vectorial de las funciones reales continuas sobre $[\alpha, \beta]$ de **R**, demostrar que las aplicaciones siguientes de E en $\mathbf{R}_+$ son normas

$$x \to \int_{\alpha}^{\beta} |x(t)|\, dt, \qquad x \to \sup_{\alpha \leq x \leq \beta} |x(t)|.$$

b) Norma asociada a una forma cuadrática no degenerada positiva

Sea f una forma bilineal simétrica *positiva* sobre E y q la forma cuadrática asociada, para todo x de E y todo λ de **R** se tiene

$$q(\lambda x) = \lambda^2 q(x) \Rightarrow \sqrt{q(\lambda x)} = |\lambda| \sqrt{q(x)}$$

por otra parte, para todo par (x, y) se tiene (ver § 230, cor. 1 del t. 14)

$$\sqrt{q(x + y)} \leq \sqrt{q(x)} + \sqrt{q(y)}.$$

Supongamos, además, f *no degenerada,* como f es positiva el único vector isótropo es 0 (ver § 230, cor. 2, del t. 14), luego

de donde: $$q(x) = 0 \Leftrightarrow x = 0$$

Teorema 16. — *Si* q *es la forma cuadrática asociada a una forma bilineal simétrica no degenerada positiva sobre* E, *espacio vectorial sobre* **R**, *la aplicación de* E *en* $\mathbf{R}_+$, *definida por*

$$x \to \sqrt{q(x)}$$

es una norma sobre E.

Definición 9. — E *espacio vectorial sobre* **R**, *provisto de una norma* $x \to \sqrt{q(x)}$, *siendo* q *una forma cuadrática no degenerada positiva sobre* E, *se llama* espacio prehilbertiano real *en el caso general y* espacio euclídeo *si es de dimensión finita.*

OBSERVACION

Algunos autores reservan el término «espacio euclídeo» al espacio afín (ver *Geometría*) asociado a un espacio vectorial de dimensión finita provisto de la norma anterior; así en esta terminología «el espacio ordinario» es el espacio euclídeo de dimensión 3.

Para evitar confusiones se puede emplear los dos términos «espacio vectorial euclídeo» y «espacio afín euclídeo». De acuerdo con la definición que hemos adoptado anteriormente, en lo sucesivo *«espacio euclídeo»* será sinónimo de *«espacio vectorial euclídeo»*.

EJEMPLOS Y EJERCICIOS

5. Demostrar que en $E = \mathbf{R}^n$ descrito por $x = (x^1, \ldots, x^4)$, la aplicación de E en $\mathbf{R}_+$ definida por

$$x \to \sqrt{(x^1)^2 + \ldots + (x^n)^2}$$

es una norma sobre E. (El caso de $n = 2$ ha sido estudiado en el § 118.)

6. Demostrar que

$$x \to \left[\int_\alpha^\beta [x(t)]^2 \, dt \right]^{1/2}$$

es una norma sobre E, espacio vectorial de las funciones reales continuas, sobre $[\alpha, \beta]$ (ver § 230, ej. 2).

c) Espacio euclídeo de dimensión n

Sea E un *"espacio euclídeo de dimensión" n, es decir, un espacio vectorial de dimensión n sobre* **R**, *provisto de una forma bilineal simétrica no degenerada positiva.*

Cualquiera que sea esta forma f, respecto a una base ortonormal relativamente a f, se tendrá

$$f(x, y) = \sum_{i=1}^{n} x^i y^i, \quad q(x) = f(x, x) = \sum_{i=1}^{n} (x^i)^2.$$

Esto nos permite decir: *Existe una sola estructura de espacio euclídeo de dimensión n.* Para estudiarlo escogeremos una forma bilineal simétrica no degenerada positiva f_0 que se llamara *forma fundamental del espacio euclídeo* E. Las bases ortonormales de E serán las bases ortonormales relativamente a f_0 (o bien a q_0 forma cuadrática asociada a f_0).

f_0 se llama entonces, generalmente, *producto escalar* sobre E. Según los autores las anotaciones siguientes se utilizan para $f_0(x, y)$

$$x \cdot y, \qquad (x \mid y), \qquad \langle x, y \rangle, \qquad \langle x \mid y \rangle$$

utilizaremos la notación $x \cdot y$, que generaliza la notación del producto escalar utilizado en el espacio ordinario, que es como hemos visto (observación anterior) el espacio afin asociado al espacio vectorial euclídeo de tres dimensiones. Guardaremos la notación $(x \mid y)$ para el producto hermitiano (ver § 237) y la notación $\langle x, y \rangle$ para el caso en que queramos distinguir $x \in \mathrm{E}$ e $y \in \mathrm{E}^*$.

$f_0(x, x) = x \cdot x$ se llama *cuadro escalar* de x: es el cuadrado de la norma $\| x \|$ (asociada a f_0), llamada *norma euclídea* de x, se dice también que $\| x \|$ es la *longitud* de x, y se le designa igualmente $| x |$.

Un vector de longitud 1 se llama vector unitario. En una base ortonormal cualquiera relativamente a f_0, se tiene

$$x \cdot y = x^1y^1 + \ldots + x^ny^n, \qquad \| x \| = \sqrt{(x^1)^2 + \ldots + (x^n)^2}.$$
$$[(\forall\, y \in \mathrm{E}) \qquad x \,.\, y = x' \cdot y] \Rightarrow x = x'.$$

Al ser la forma f_0 no degenerada positiva, el único vector isótropo de un espacio euclídeo es 0, y tenemos

El *adjunto* u^* de un endomorfismo u de E estará entonces definido con ayuda del producto escalar por

$$[\forall\, (x, y) \in \mathrm{E}^2] \qquad u(x) \cdot y = x \,.\, u^*(y)$$

un operador *ortogonal* por

$$[\forall\, (x, y) \in \mathrm{E}^2] \qquad u(x) \cdot u(y) = x \,.\, y$$

y un operador *simétrico* por

$$[\forall\, (x, y) \in \mathrm{E}^2] \qquad u(x) \cdot y = x \,.\, u(y)$$

todos estos calificativos —adjunto, ortogonal, simétrico— son relativos a la forma fundamental escogida.

El grupo ortogonal $\mathbf{O}(f_0)$ es isomorfo al grupo $\mathbf{O}(n, \mathbf{R})$ de las matrices reales ortogonales de orden n, lo designaremos por esta notación $\mathbf{O}(n, \mathbf{R})$. Designaremos $\mathbf{O}^+(n, \mathbf{R})$, $\mathbf{SO}(n, \mathbf{R})$ y el *grupo de las rotaciones* y el *grupo de las matrices ortogonales de orden n de determinante* $+ 1$ (ver § 228, *b*). El espacio euclídeo de n dimensiones cuando es *real* podemos *orientarlo* (ver § 172). Lo que hemos dicho de este § 172 y del resultado del § 228 nos permiten enunciar:

En un espacio euclídeo de dimensión n toda matriz de paso de una base ortonormal de igual orientación *es una matriz ortogonal de determinante* $+ 1$.

Sea, en fin, una forma bilineal simétrica f sobre E, degenerada o no, positiva o no, en una base ortonormal; tendremos con las notaciones acostumbradas

$$f(x, y) = {}^t\mathrm{Y}\mathrm{A}\mathrm{X} \qquad ({}^t\mathrm{A} = \mathrm{A}).$$

Las consideraciones del final del § 229 permiten asociarle de manera biyectiva un endomorfismo simétrico u, aquel cuya matriz respecto a una base considerada es A, y tenemos

$$[\forall (x, y) \in \mathrm{E}^2] \qquad f(x, y) = u(x) \cdot y = x \cdot u(y).$$

Por lo tanto, estudiar u o f (o bien q forma cuadrática asociada a f) o también A matriz cuadrada real simétrica de orden n, es lo mismo: y es lo que vamos hacer en el párrafo siguiente.

EJERCICIOS

7. Demostrar que toda familia (x_i) de vectores no nulos, ortogonales dos a dos, de un espacio euclídeo, es libre.

8. Si (a_i) $(1 \leq i \leq n)$ es una base *cualquiera* del espacio euclídeo E, demostrar que hay una sola base ortogonal (b_i), cuyos vectores son de la forma

$$\begin{aligned} b_1 &= a_1 \\ b_2 &= a_2 + \lambda_2^1 b_1 \\ &\vdots \\ b_i &= a_i + \lambda_i^1 b_1 + \dots + \lambda_i^{i-1} b_{i-1} \\ &\vdots \\ b_n &= a_n + \lambda_n^1 b_1 + \dots + \lambda_n^{n-1} b_{n-1} \end{aligned}$$

(se escribirá $b_i \cdot b_1 = \dots = b_i \cdot b_{i-1} = 0$ para $i = 2, \dots, n$, y se utilizará el ejercicio precedente).

9. Deducir del ejercicio anterior que dada una base *cualquiera* (a_i) del espacio euclídeo E, de dimensión n, existe una base *ortonormal única* (c_i) de E, tal que para todo entero k de $[1, n]$:

a) El subespacio engendrado por $a_1, \dots, a_k$ sea idéntico al subespacio engendrado por $c_1, \dots, c_k$.

b) $a_k \cdot c_k > 0$.

Este procedimiento se le conoce con el nombre de *ortonormalización de Schmitt*.

10. Si E es un espacio euclídeo *orientado*, de dimensión n, F es un hiperplano de E, y $(a_1, a_2, \dots, a_{n-1})$ una base ortonormal de F. Demostrar que existe una única base ortonormal directa de E de la forma $(a_1, a_2, \dots, a_{n-1}, a_n)$. (Observar que dim $\mathrm{F}^\perp = 1$.)

11. Si (x_i) $(1 \leq i \leq m)$, $m \leq n$ es una familia de vectores ortogonales, dos a dos, de un espacio vectorial de dimensión n. Demostrar que (teorema de PITÁGORAS)

$$\| x_1 + \dots + x_m \|^2 = \| x_1 \|^2 + \dots + \| x_m \|^2.$$

232. Complexificación de un espacio vectorial sobre R. Aplicaciones

a) Complexificado de un espacio vectorial sobre R

Si E es un espacio vectorial de dimensión n sobre **R** y (a_j) una base de E, se tiene

$$E = \mathbf{R}a_1 \oplus \ldots \oplus \mathbf{R}a_n$$

consideremos el espacio vectorial E′ sobre **C**, que tiene igualmente por base $(a_i)\,(1 \leq i \leq n)$, se tendrá

$$E' = \mathbf{C}a_1 \oplus \ldots \oplus \mathbf{C}a_n.$$

Si se toma otra base (b_i) en E, está claro que $E'' = \mathbf{C}b_1 \oplus \ldots \oplus \mathbf{C}b_n$ es idéntico a E′ cuando las b_i son combinaciones lineales con coeficientes reales "de" las a_i y recíprocamente. *El espacio E′, espacio vectorial de dimensión n sobre* **C**, *asociado así a* E, *espacio vectorial de dimensión n sobre* **R**, *se llama "complexificado" de* E, y se tiene

$$E \subset E', \qquad E' \neq E.$$

EJERCICIO

1. Demostrar que E′ se puede considerar como un espacio vectorial sobre **R**, y que es entonces de dimensión $2n$. (Se demostrará que $\{a_1, \ldots, a_n, ia_1, \ldots, ia_n\}$ es una base de E′, espacio vectorial sobre **R**.) Deducir de lo anterior que E es un subespacio de E′, espacio vectorial sobre **R**.

Si se supone E *euclídeo,* es decir, provisto de una forma bilineal simétrica no degenerada positiva f_0

$$f_0\colon\ E \times E \rightarrow \mathbf{R}$$

se tendrá, cuando (a_i) es una base ortonormal relativamente a f_0

(1) $$[\forall\,(x, y) \in E^2] \qquad f_0(x, y) = \sum_{i=1}^{n} x^i y^i.$$

Se puede entonces dar a E′ la forma bilineal f_0'

$$f_0'\colon\ E' \times E' \rightarrow \mathbf{C}$$

definida en la base (a_i) (base ortonormal de E respecto a f_0) por

(1) $$[\forall\,(x, y) \in E'^2] \qquad f_0'(x, y) = \sum_{i=1}^{n} x^i y^i$$

con $x = \sum_{i=1}^{n} x^i a_i$, $y = \sum_{i=1}^{n} y^i a_i$, pero en este último caso $x^1, \ldots, x^n, y^1, \ldots, y^n$

son *complejos.* La fórmula (1′) demuestra que f'_0 es una forma bilineal simétrica no degenerada sobre E′ y que (a_i) considerada como base de E′ es ortonormal respecto a f'_0. Pero en este caso

$$f'_0(x, x) = \sum_{i=1}^{n} (x^i)^2$$

es un *número complejo que puede ser nulo* $x \neq 0$ (ver 225, ej. 3), hay en E′ vectores isótropos no nulos. No se puede asociar norma a f'_0, y f'_0 *no es un producto escalar.*

b) Endomorfismos asociados en E y en el complexificado E′ de E

Cuando u es un endomorfismo de E se le puede asociar un endomorfismo u' de E′, definido por

$$M[u', (a_i)] = M[u, (a_i)] = A$$

las fórmulas de cambio de bases en E demuestran que la definición de u' es independiente de la base (a_i) común a E y E′ escogida. *Diremos que u' endomorfismo de E′ está asociado al endomorfismo u de* E. Los valores propios de u (en **C**) son los mismos que los de u', que son los de la matriz real A. En particular si se supone u ortogonal (resp. simétrico) respecto a f_0, u' será ortogonal (resp. simétrico) relativamente a f'_0, puesto que las matrices de u y u', en una base (a_i) ortonormal relativamente a f_0 y f'_0, al ser las mismas, son al mismo tiempo ortogonales o simétricas.

Estudiemos los *endomorfismos ortogonales* de E′ se tendrá

$$(2) \qquad [\forall (x, y) \in E'^2] \qquad f'_0[u'(x), u'(y)] = f'_0(x, y)$$

sea λ un valor propio de u' (por tanto de A y de u en **C**) y x un vector propio asociado *no nulo,* se tendrá

$$f'_0(\lambda x, \lambda x) = \lambda^2 f'_0(x, x) = f'_0(x, x)$$

si λ es real, x pertenece a E y al ser no nulo $f'_0(x, x) = f_0(x, x) \neq 0$, luego $\lambda^2 = 1$: *los únicos valores propios reales de un operador ortogonal de un espacio euclídeo son, por tanto,* ± 1 (ver § 228, ej. 3).

Si λ no es real, se tendrá $\lambda^2 \neq 1$ los vectores propios asociados no nulos no pertenecen a E, la relación anterior demuestra que $f'_0(x, x) = 0$, en consecuencia:

Si u es un operador ortogonal del espacio euclídeo E, *los vectores propios de u', asociado a u, relativos a los valores propios no reales de u' (es decir, no reales de u) son vectores isótropos del complexificado* E′ *de* E.

EJERCICIOS

2. Demostrar que todo valor propio de u', operador de E′, asociado a u, operador ortogonal del espacio euclídeo E, no real, es de módulo 1 (aplicar la fórmula (2)

$$\text{a } x = \sum_{i=1}^{n} x^i a_i \neq 0 \quad \text{asociado a } \lambda \text{ y a } \quad y = \bar{x} = \sum_{i=1}^{n} \bar{x}^i a_i \quad \text{asociado a } \bar{\lambda}).$$

(En el ejercicio 2 del § 239 hay otra demostración más elegante.)

3. Demostrar que si E es un espacio euclídeo de dimensión *impar*, toda rotación de E tiene un número impar de valores propios iguales a $+1$, por lo tanto, al menos una. (Observar que el producto de los valores propios de u en **C** es $\det u = 1$, ver § 218, *b*) y § 209.)

4. Demostrar que toda matriz de permutación (ver capítulo 8, ej. 204) es ortogonal. ¿Qué ocurre con las matrices correspondientes a una permutación par?

Vamos ahora en parte gracias al complexificado del espacio euclídeo E, a estudiar los endomorfismos simétricos E.

233. Diagonalización de un operador simétrico real

Sea u un operador simétrico del espacio euclídeo E de dimensión n sobre **R**, A su matriz respecto a una base ortonormal de E, nos proponemos demostrar que A, por tanto u, tienen todos sus valores propios reales y que son siempre diagonalizables.

a) Propiedades de los valores y de los vectores propios de un operador simétrico real

Sea E′ el complexificado de E y u' el operador de E′ asociado al operador simétrico u de E con las notaciones del párrafo precedente, tendremos

$$[\forall (x, y) \in E'^2] \qquad f'_0[u'(x), y] = f'_0[x, u'(y)].$$

Sea λ un valor propio de u' (por tanto de u y de A) en **C**. Si λ es complejo, u' admite el valor propio $\bar{\lambda}$, sea x y $\bar{x}$ dos vectores propios, no nulos, asociados, respectivamente, a λ y $\bar{\lambda}$ y teniendo sus coordenadas dos a dos, conjugadas en una base (a'_i) ortonormal relativamente a f_0 y f'_0. La relación anterior nos da

$$f'_0(\lambda x, \bar{x}) = f'_0(x, \bar{\lambda}\bar{x})$$

de donde

$$(\lambda - \bar{\lambda}) f'_0(x, \bar{x}) = (\lambda - \bar{\lambda}) \sum_{i=1}^{n} |x^i|^2 = 0$$

$x \neq 0$ implicando $\sum_{i=1}^{n} |x^i|^2 \neq 0$, $\lambda = \bar{\lambda}$ por tanto es real.

La demostración anterior es pesada (recurre al complexificado de E); daremos en el § 240, *c* una demostración idéntica a esta, pero más elegante, utilizando el producto hermitiano

Los n valores propios de un operador simétrico de E espacio vectorial de dimensión n al ser reales, a partir de ahora no tendremos más necesidad del complexificado E′ de E.

Sea x un vector propio no nulo de u, asociado a λ, tenemos

$$x \cdot y = 0 \Rightarrow x \cdot u(y) = u(x) \cdot y = \lambda (x \cdot y) = 0$$

por tanto si y es ortogonal a x vector propio no nulo de u, también lo es de $u(y)$, dicho de otra manera el *subespacio ortogonal a un vector propio no nulo de u es estable por u.*

Sea finalmente dos vectores propios de u, x_1 y x_2 asociados, respectivamente, a λ_1 y λ_2 supuestos *distintos,* tendremos

$$u(x_1) \cdot x_2 = x_1 \cdot u(x_2)$$

es decir,

$$\lambda_1(x_1 \cdot x_2) = \lambda_2(x_1 \cdot x_2) \Rightarrow (\lambda_1 - \lambda_2)(x_1 \cdot x_2) = 0$$

por tanto x_1 y x_2 son ortogonales, de lo anterior resulta que los subespacios propios $V(\lambda_1)$ y $V(\lambda_2)$ (ver § 217) son ortogonales. Resumamos estas propiedades:

TEOREMA 17. — *Si u es un endomorfismo simétrico del espacio euclídeo* E *de dimensión n:*

1. *Los n valores propios de u, son reales.*
2. *El subespacio ortogonal a todo vector propio no nulo de u, es estable por u.*
3. *Los subespacios propios asociados a dos valores propios distintos, son ortogonales.*

b) Diagonalización de un endomorfismo simétrico real

Nos proponemos demostrar que todo endomorfismo simétrico u de un espacio euclídeo de dimensión n es diagonalizable; de un modo más preciso vamos a demostrar que existen bases ortonormales de E, formadas de vectores propios de u. Vamos a utilizar el mismo método que en el § 220 (reducción a la forma triangular de una matriz cuadrada compleja). Pero aquí la estabilidad por u del subespacio ortogonal a un vector propio de u simplificará el resultado.

Sea u un endomorfismo simétrico de E, y A su matriz respecto a una base ortonormal (a_i) de E.

Si $n = 1$, u tiene un solo valor propio λ, le corresponde $x \neq 0$, vector propio, $b = \dfrac{x}{\|x\|}$ constituye una base de E, b es un vector propio, y, además, $\|b\| = 1$.

Sea $n > 1$, supongamos que para todo endomorfismo simétrico de un espacio vectorial euclídeo de dimensión $n - 1$ existe una base ortonormal de vectores propios (hipótesis de recurrencia), u endomorfismo simétrico de E de dimensión n tiene n valores propios reales. A una de ellas λ_1 se puede asociar un vector propio b_1, tal que $\|b_1\| = 1$. Consideremos el ortogonal $F^\perp$ de $F = \mathbf{R}b_1$, F no es isótropo puesto que en E euclídeo el único vector isótropo es 0, por tanto

$$E = F \oplus F^\perp.$$

La restricción de la forma fundamental f_0 de E a $F^\perp$ es siempre no degenerada positiva, existen por tanto bases de $F^\perp$ ortonormales; para cada una de ellas $(b'_2, \ldots, b'_n)$ es claro que $(b_1, b'_2, \ldots, b'_n)$ es una base ortonormal de E,

$$B' = \left(\begin{array}{c|c} \lambda_1 & 0 \ \dots \ 0 \\ \hline 0 & \\ \vdots & C' \\ 0 & \end{array}\right)$$

pues $F^\perp$ *es estable por* u. El bloque C' (matriz cuadrada de orden $n-1$) es $M[u', (b'_i)(2 \leq i \leq n)$, al ser u' el endomorfismo inducido por u en $F^\perp$. u' es simétrico, en efecto

$$[\forall (x, y) \in F^\perp] \qquad u'(x) \cdot y = u(x) \cdot y = x \cdot u(y) = x \cdot u'(y).$$

Por otra parte si x es un vector propio de u' existe λ' real, tal que

$$u'(x) = \lambda' x \Rightarrow u(x) = u'(x) = \lambda' x$$

por tanto los vectores propios de u' son los vectores propios de u. Aplicando la hipótesis de recurrencia a u' vemos que existe para $F^\perp$ una base ortonormal $(b_2, \dots, b_n)$ de vectores propios de u' por tanto de u y en consecuencia $(b_1, b_2, \dots, b_n)$ es una base ortonormal de E formada de vectores propios de u. La propiedad al ser cierta para $n = 1$, también es cierta para todo n:

TEOREMA 18. — *Si* u *es un endomorfismo simétrico de* E, *espacio euclídeo de dimensión* n, *existen bases ortonormales de* E *formadas por vectores propios de* u.

Sea B la matriz de u en la base $(b_i)(1 \leq i \leq n)$, es diagonal y los elementos diagonales son los valores propios $\lambda_1, \dots, \lambda_n$ de u, asosiados, respectivamente, a los elementos de la base $(b_1, \dots, b_n)$. Las dos bases (a_i) y (b_i) al ser ortonormales, la matriz de paso P es ortogonal (ver § 228, *a*), por tanto:

COROLARIO 1. — *Para toda matriz simétrica real* A, *de orden* n, *existe una matriz* P *ortogonal real, de orden* n, *tal que la matriz* $B = P^{-1}AP$ *sea diagonal.*

c) Reducción de las formas cuadráticas reales

Sea finalmente una forma bilineal simétrica cualquiera f sobre E, espacio euclídeo de dimensión n; hemos visto al final del § 231 que se le puede asociar de manera biyectiva un endomorfismo simétrico u de E, tal que

$$[\forall (x, y) \in E^2] \qquad f(x, y) = u(x) \cdot y = {}^t(AX)Y = {}^tXAY$$

Si A es la matriz asociada a f, o a u, en una base ortonormal cualquiera. En una base ortonormal formada de vectores propios asociados, respectivamente, a los valores propios $\lambda_1, \dots, \lambda_n$ de A, se tendrá:

$$f(x, y) = \sum_{i=1}^{n} \lambda_i x^i y^i, \qquad q(x) = f(x, x) = \sum_{i=1}^{n} \lambda_i (x^i)^2.$$

Naturalmente si rg $(f) = r$, cambiando si fuera preciso la numeración de la base, se tendrá

$$f(x, y) = \sum_{i=1}^{r} \lambda_i x^i y^i, \qquad q(x) = f(x, x) = \sum_{i=1}^{r} \lambda_i (x^i)^2$$

donde los $\lambda_i (1 \leq i \leq r)$ son no nulos.

La operación anterior se llama la *"reducción de las formas cuadráticas"* en una base ortonormal (respecto a la forma fundamental f_0 de E).

La base (b_i) considerada anteriormente es por tanto ortonormal relativamente a la forma fundamental f_0 y ortogonal relativamente a la forma f, de donde:

COROLARIO 2. — *Si* E *es un espacio vectorial de dimensión n sobre* **R**, f_0 *una forma bilineal simétrica no degenerada positiva sobre* E, *y f una forma bilineal simétrica cualquiera sobre* E, *existen bases de* E *ortonormales relativamente a f, y ortogonales relativamente a f.*

EJERCICIOS

1. Designemos por $\lambda_1, \ldots, \lambda_m$ los $m(m \leq n)$ valores propios distintos dos a dos de u endomorfismo simétrico de E y por $V(\lambda_1), \ldots, V(\lambda_m)$ los subespacios propios asociados. Demostrar el teorema 18, demostrando que

$$E = V(\lambda_1) \oplus \ldots \oplus V(\lambda_m).$$

2. En $E = \mathbf{R}^3$ referido a su base canónica se considera la forma bilineal asociada a la forma cuadrática definida por

$$q(x) = -[(x^1)^2 + (x^2)^2 + (x^3)^2] + 2[x^2x^3 + x^3x^1 + x^1x^2]$$

demostrar que existe una base ortonormal de E, tal que

$$q(x) = (x'^1)^2 - 2[(x'^2)^2 + (x'^3)^2]$$

(ver § 219, ejercicio 3).

234. Producto mixto y producto vectorial en el espacio euclídeo orientado de dimensión 3

En todo este párrafo E designa el espacio vectorial euclídeo orientado de dimensión 3.

a) Producto mixto

Si (a_i) es una base cualquiera del espacio vectorial E, existe una forma bilineal alternada única que toma el valor 1 para (a_1, a_2, a_3) (ver § 167, *a*), ésta es la aplicación

$$(x_1, x_2, x_3) \rightarrow \det{}_{(a_i)}(x_1, x_2, x_3)$$

esta aplicación de E × E × E en **R** no es intrínseca: depende de la base (a_i) escogida. Pero supongamos que (a_i) y (a'_i) *sean bases ortonormales directas.*

Si x_1, x_2, x_3 son tres vectores cualesquiera de E, consideremos los tres endomorfismos de E, u, f, f', definidos de la manera siguiente (u es un automorfismo)

$$(i = 1, 2, 3) \qquad u(a'_i) = a_i, \qquad f(a_i) = x_i, \qquad f'(a'_i) = x_i$$

tenemos

$$f' = f \circ u \Rightarrow \det f' = (\det f)(\det u)$$

ahora bien,

$$\det f = \det_{(a_i)}(x_1, x_2, x_3), \qquad \det f' = \det_{(a_i')}(x_1, x_2, x_3)$$

y $\det u = 1$, pues u es un endomorfismo ortogonal con determinante positivo (ver § 231, *c*), de donde:

Teorema y definición. — *La aplicación de* E × E × E *en* **R** *definida por*

(1) $$(x_1, x_2, x_3) \rightarrow \det_{(a_i)}(x_1, x_2, x_3)$$

es independiente de la base ortonormal directa (a_i) *se le llama* producto mixto.

Por abuso de lenguaje se llama *producto mixto* igualmente el número real $\det_{(a_i)}(x_1, x_2, x_3)$ y se le representa (x_1, x_2, x_3), el contexto indica si se trata de un elemento de E × E × E o del número real $\det_{(a_i)}(x_1, x_2, x_3)$. Las *propiedades del producto mixto* siguen las propiedades de los determinantes:

1. El producto mixto es *lineal* respecto a cada vector, es *alternado* por tanto *antisimétrico* (la característica de **R** es 0), por tanto

$$(x_1, x_2, x_3) = (x_2, x_3, x_1) = (x_3, x_1, x_2)$$
$$= -(x_2, x_1, x_3) = -(x_1, x_3, x_2) = -(x_3, x_2, x_1)$$

pues una permutación circular que opera sobre un conjunto con tres elementos es par.

2. $(x_1, x_2, x_3) \neq 0$ si y solamente si la familia (x_i) es libre.

En particular $(x_1, x_2, x_3) > 0$ si y solamente si (x_i) es una base directa del espacio vectorial E.

3. Finalmente en toda base ortonormal directa (a_i)

$$(x_1, x_2, x_3) = \begin{vmatrix} x_1^1 & x_2^1 & x_3^1 \\ x_1^2 & x_2^2 & x_3^2 \\ x_1^3 & x_2^3 & x_3^3 \end{vmatrix}$$

si x_i^j es la coordenada de x_i sobre a .

b) **Producto vectorial**

Si x_1 y x_2 son dos vectores cualesquiera de E, consideremos la aplicación de en **R**, definida por

$$y \rightarrow (x_1, x_2, y) \in \mathbf{R}$$

es una forma lineal sobre E, es decir, un elemento z^* de E^*, por tanto

$$(x_1, x_2, y) = z^*(y) = \langle y, z^* \rangle$$

si se ha identificado E y E^* por el isomorfismo canónico φ (asociado a la forma fundamental f_0 de E) tendremos poniendo $z^* = z \in E$

$$(x_1, x_2, y) = \langle y, z^* \rangle = f_0(y, z) = y \cdot z = z \cdot y.$$

TEOREMA Y DEFINICIÓN. — *En* E *espacio vectorial euclídeo orientado de dimensión* 3, *para todo par* (x_1, x_2) *de vectores* E *existe un vector único* z, *tal que*

(2) $$(\forall y \in E) \qquad (x_1, x_2, y) = z \cdot y = y \cdot z$$

este vector se llama producto vectorial *de* x_1 *y* x_2 *y se representa* $x_1 \wedge x_2$.

Se tiene, por tanto

$$(x_1, x_2, x_3) = (x_1 \wedge x_2) \cdot x_3$$

que es el origen de la expresión "*producto mixto*": es el *producto escalar* de un *producto vectorial* y de un vector.

Las propiedades del producto vectorial proceden de la relación (2).

1. Si x_1 y x_2 son dependientes $(x_1, x_2, y) = (x_1 \wedge x_2) \cdot y = 0$ para todo y por tanto $x_1 \wedge x_2 = 0$. Si $x_1 \wedge x_2 = 0$ para todo y, $(x_1, x_2, y) = 0$. En consecuencia si x_1 y x_2 son independientes, existe y, tal que $(x_1, x_2, y) \neq 0$ y $x_1 \wedge x_2 \neq 0$. De donde: $x_1 \wedge x_2 = 0$ *si y sólo si* x_1 *y* x_2 *son dependientes*. En particular para todo x, $x \wedge x = 0$.

2. La aplicación $E \times E$ en E (de hecho en E^*) definida por $(x_1, x_2) \to x_1 \wedge x_2$ es *bilineal*; en efecto,

$$(x_1 + x_1', x_2, y) = (x_1, x_2, y) + (x_1', x_2, y)$$

puede escribirse

$$\begin{aligned}[(x_1 + x_1') \wedge x_2] \cdot y &= (x_1 \wedge x_2) \cdot y + (x_1' \wedge x_2) \cdot y \\ &= [(x_1 \wedge x_2) + (x_1' \wedge x_2)] \cdot y\end{aligned}$$

y esto para todo y de donde cualesquiera que sean x_1, x_2, x_1'

$$(x_1 + x_1') \wedge x_2 = (x_1 \wedge x_2) + (x_1' \wedge x_2).$$

Se demostrará igualmente

$$\begin{aligned}x_1 \wedge (x_2 + x_2') &= (x_1 \wedge x_2) + (x_1 \wedge x_2') \\ (\lambda x_1) \wedge x_2 = x_1 \wedge (\lambda x_2) &= \lambda(x_1 \wedge x_2).\end{aligned}$$

Hemos visto más arriba que $x \wedge x = 0$, la aplicación $(x_1, x_2) \to x_1 \wedge x_2$ es, pues, *bilineal alternada*; por tanto, *antisimétrica*, pues

$$0 = (x_1 + x_2) \wedge (x_1 + x_2) = (x_1 \wedge x_1) + (x_1 \wedge x_2) + (x_2 \wedge x_1) + (x_2 \wedge x_2)$$

implica

$$x_2 \wedge x_1 = -(x_1 \wedge x_2).$$

3. $x_1 \wedge x_2$ es *ortogonal* a x_1 y x_2, pues

$$(x_1 \wedge x_2) \cdot x_1 = (x_1, x_2, x_1) = 0, \qquad (x_1 \wedge x_2) \cdot x_2 = (x_1, x_2, x_2) = 0.$$

4. Si x_1 y x_2 son independientes los tres vectores (en este orden) $x_1, x_2, x_1 \wedge x_2$ describen una *base directa* de E; en efecto,

$$(x_1, x_2, x_1 \wedge x_2) = (x_1 \wedge x_2) \cdot (x_1 \wedge x_2) = \| x_1 \wedge x_2 \|^2 > 0.$$

En particular si (a_i) es una base ortonormal directa: $a_1 \wedge a_2 = \lambda a_3$ $(\lambda \in \mathbf{R})$, pues $a_1 \wedge a_2$ es ortogonal a a_1 y a a_2; por otra parte

$$1 = (a_1, a_2, a_3) = (a_1 \wedge a_2) \cdot a_3 = \lambda \| a_3 \|^2 = \lambda$$

por tanto

$$a_1 \wedge a_2 = a_3$$

se demostrará igualmente

$$a_2 \wedge a_3 = a_1, \qquad a_3 \wedge a_1 = a_2.$$

5. Sea (a_i) una base ortonormal directa de E y dos vectores

$$x_1 = x_1^1 a_1 + x_1^2 a_2 + x_1^3 a_3, \qquad x_2 = x_2^1 a_1 + x_2^2 a_2 + x_2^3 a_3$$

las propiedades precedentes demuestran

$$x_1 \wedge x_2 = (x_1^2 x_2^3 - x_1^3 x_2^2) a_1 + (x_1^3 x_2^1 - x_1^1 x_2^3) a_2 + (x_1^1 x_2^2 - x_1^2 x_2^1) a_3$$

las coordenadas del producto vectorial $x_1 \wedge x_2$ en la base (a_i), son por tanto los adjuntos de y^1, y^2, y^3, en el determinante

$$\begin{vmatrix} x_1^1 & x_2^1 & y^1 \\ x_1^2 & x_2^2 & y^2 \\ x_1^3 & x_2^3 & y^3 \end{vmatrix}$$

OBSERVACION

Si se cambia la orientación del espacio, el producto mixto y el producto vectorial se cambian por sus opuestos.

EJERCICIOS

1. Si y y z son independientes, demostrar que existe λ y μ únicos, tales que

$$y \wedge (y \wedge z) = \lambda y + \mu z$$
$$(\lambda = y \cdot z,\ \mu = -y^2).$$

2. Demostrar la fórmula del «*doble producto vectorial*»

$$x \wedge (y \wedge z) = (x \cdot z) y - (x \cdot y) z$$

(escribir $x = \alpha y + \beta z + \gamma (y \wedge z)$ cuando y y z son independientes, y aplicar el ejercicio anterior).

3. Demostrar la fórmula del doble producto vectorial utilizando las coordenadas de los vectores.

4. Si a y b son dos vectores ortogonales ($a \neq 0$). Demostrar que las soluciones de la ecuación

$$a \wedge x = b$$

son de la forma $x = x_0 + \lambda a$ (λ número real cualquiera). Demostrar que existe x único, tal que $a \cdot x = 0$. (Comparar este ejercicio con el ejercicio 2 del § 177).

235. Angulos de dos vectores en el espacio euclídeo orientado de dimensión 2 o 3

Representaremos por E_2 (resp. E_3) el espacio euclídeo orientado de dimensión 2 (resp. 3). Representamos los vectores por las letras latinas con una flecha encima. Recordemos que (ver § 231, ej. 10) dado $\vec{a}$ unitario, existe en E_2 una sola base $(\vec{a}, \vec{b})$ ortonormal directa y que dada una base ortonormal $(\vec{a}, \vec{b})$ de un plano F de E_3 existe una sola base $(\vec{a}, \vec{b}, \vec{c})$ ortonormal directa de E_3.

a) Isomorfismo entre los grupos $O^+(2, \mathbf{R})$ y U

Busquemos las matrices reales ortogonales de determinante $+1$. Las relaciones

$$A = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}, \qquad {}^tAA = I^2, \qquad \det A = +1$$

implican

$$\alpha^2 + \beta^2 = 1, \qquad \alpha\gamma + \beta\delta = 0, \qquad \gamma^2 + \delta^2 = 1, \qquad \alpha\delta - \beta\gamma = 1$$

donde α, β no son nulos simultáneamente, la segunda ecuación da

$$\frac{\gamma}{-\beta} = \frac{\delta}{\alpha} = \lambda$$

la cuarta ecuación da $\lambda = 1$, de donde $\gamma = -\beta$, $\delta = \alpha$ y la tercera ecuación se verifica si y solamente si la primera también se verifica, por tanto

$$A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \qquad \text{con} \qquad \alpha^2 + \beta^2 = 1$$

pues si $u = \alpha + i\beta \in \mathbf{U}$ (ver § 116) se puede definir la aplicación $u \to \mathbf{A}$ que es claramente biyectiva. Por otra parte

$$\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}\begin{pmatrix} \alpha' & -\beta' \\ \beta' & \alpha' \end{pmatrix} = \begin{pmatrix} \alpha\alpha' - \beta\beta' & -(\alpha\beta' + \beta\alpha') \\ \alpha\beta' + \beta\alpha' & \alpha\alpha' - \beta\beta' \end{pmatrix}$$

de donde:

LEMA 1. — *La aplicación del grupo multiplicativo* U *de los números complejos de módulo* 1 *en el grupo de las matrices de tipo* (2,2) *reales ortogonales de determinante* + 1 *definido por*

$$u = \alpha + i\beta \to \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$$

es un isomorfismo de grupos.

El grupo de las matrices (2, 2) reales ortogonales de determinante 1 es por tanto *conmutativo.*

Sea r una rotación de E_2, es decir, un elemento de $\mathbf{O}^+(2, \mathbf{R})$. En una base ortonormal directa tiene asociada una matriz del tipo anterior. Sean $(\vec{a}, \vec{b})$ y $(\vec{a}' \ \vec{b}')$ dos bases ortonormales directas, tenemos

$$M[r, (\vec{a}', \vec{b}')] = P^{-1}M[r, (\vec{a}, \vec{b})]P = M[r, (\vec{a}, \vec{b})]$$

pues si **P** es la matriz de paso de la primera base a la segunda, las cuatro matrices consideradas en la igualdad anterior son ortogonales de determinante + 1, de donde

LEMA 2. — *La aplicación* f_1 *del grupo* U *en el grupo* $\mathbf{O}^+(2, \mathbf{R})$ *de las rotaciones de* E_2 *definido por*

$$u = \alpha + i\beta \in U \qquad r = f_1(u) \qquad r(\vec{a}) = \alpha\vec{a} + \beta\vec{b}$$

es independiente de la base ortonormal directa $(\vec{a}, \vec{b})$ *escogida,* f_1 *es un isomorfismo de grupos.*

Se observará que la relación $r(\vec{a}) = \alpha\vec{a} + \beta\vec{b}$ es suficiente para definir r pues la forma de la matriz $A = M(r)$ demuestra que

$$r(\vec{b}) = -\beta\vec{a} + \alpha\vec{b}.$$

Así $r_0 = \mathrm{id}_E$ está asociada a $u = 1$, r_1 definida por $r_1(\vec{a}) = \vec{b}$ está asociada a i de donde

$$r_2 = r_1 \circ r_1 = (r_1)^2 = -\mathrm{id}_E, \qquad r_4 = r_2 \circ r_2 = (r_2)^2 = \mathrm{id}_E$$

b) Angulos de dos vectores unitarios de E_2

Sea $(\vec{u}_1, \vec{u}_2)$ un par de vectores unitarios de E_2, según una observación hecha anteriormente, existe una base ortonormal directa única $(\vec{u}_1, \vec{v}_1)$, en esta base existe una pareja única de reales tales que

$$\vec{u}_2 = \alpha\vec{u}_1 + \beta\vec{v}_1$$

existe por tanto una rotación única de E_2 tal que $r(\vec{u}_1) = \vec{u}_2$ es la rotación definida por $r(\vec{u}_1) = \alpha\vec{u}_1 + \beta\vec{v}_1$, es decir, la rotación asociada a $u = \alpha + i\beta(|u| = 1)$.

LEMA 3. — *Si* $(\vec{u}_1, \vec{u}_2)$ *son un par de vectores unitarios de* E *existe una rotación única de* E_2, *tal que* $r(\vec{u}_1) = \vec{u}_2$.

Sea $\mathcal{U}$ el conjunto de los vectores unitarios de E_2, consideremos la rotación $\mathcal{R}$ definida en $\mathcal{U} \times \mathcal{U}$ por

$$(1) \qquad [\exists r \in \mathbf{O}^+(2. \mathbf{R})] \qquad r(\vec{u}_1) = \vec{u}_2, \qquad r(\vec{u}'_1) = \vec{u}'_2$$

esta relación es evidentemente reflexiva y simétrica; es transitiva pues (1) y

$$(1') \qquad [\exists s \in \mathbf{O}^+(2, \mathbf{R})] \qquad s(\vec{u}'_1) = \vec{u}'_2, \qquad s(\vec{u}''_1) = \vec{u}''_2$$

implican, según el lema 3, $s = r$ en consecuencia

$$[\exists r \in \mathbf{O}^+(2, \mathbf{R})] \qquad r(\vec{u}_1) = \vec{u}_2, \qquad r(\vec{u}''_1) = \vec{u}''_2$$

$\mathcal{R}$ es por tanto una equivalencia lo que nos permite establecer la definición

DEFINICIÓN. — *En el plano euclídeo orientado se llama* ángulo de dos vectores unitarios $\vec{u}_1\ \vec{u}_2$, *y se representa* $\left(\widehat{\vec{u}_1, \vec{u}_2}\right)$ *toda clase de equivalencia determinada en* $\mathcal{U} \times \mathcal{U}$ *por la relación* $\mathcal{R}$. *El conjunto de los ángulos* $(\mathcal{U} \times \mathcal{U})/\mathcal{R}$ *se representará* $\mathcal{A}$.

Por tanto a toda rotación r está asociado de manera biyectiva un ángulo $\hat{\theta}$, definido por

$$r(\vec{u}_1) = \vec{u}_2 \qquad \hat{\theta} = \left(\widehat{\vec{u}_1, \vec{u}_2}\right)$$

donde el vector $\vec{u}_1$ es arbitrario; designemos por f_2 esta aplicación

$$f_2 : \mathbf{O}^+(2, \mathbf{R}) \to \mathcal{A}$$

al ser f_2 biyectivo podemos *transportar* a $\mathcal{A}$ la estructura de *grupo abeliano* de $\mathbf{O}^+(2, \mathbf{R})$ (V. § 55); al expresar aditivamente la ley así definida en $\mathcal{A}$ tendremos

$$f_2(r) + f_2(r') = f_2(r \circ r').$$

Sea $\hat{\theta} = \left(\widehat{\vec{u}_1, \vec{u}_2}\right) = f_2(r)$ y $\hat{\theta}' = \left(\widehat{\vec{u}'_1, \vec{u}'_2}\right) = f_2(r')$, según una observación hecha anteriormente existe $\vec{u}_3$ único, tal que $\hat{\theta}' = \left(\widehat{\vec{u}_2, \vec{u}_3}\right)$ se tendrá por tanto

$$\left(\widehat{\vec{u}_1, \vec{u}_2}\right) + \left(\widehat{\vec{u}_2, \vec{u}_3}\right) = \left(\widehat{\vec{u}_1, \vec{u}_3}\right)$$

pues $r \circ r' = r' \circ r$ y

$$(r' \circ r)(\vec{u}_1) = r'[r(\vec{u}_1)] = r'(\vec{u}_2) = \vec{u}_3$$

de lo anterior resulta

$$(\widehat{\vec{u}, \vec{u}}) = \hat{0}, \qquad (\widehat{\vec{u}_2, \vec{u}_1}) = -(\widehat{\vec{u}_1, \vec{u}_2})$$

se escribirá 0 en lugar de $\hat{0}$ es el *ángulo nulo.*

El diagrama

$$U \xrightarrow{f_1} O^+(2, \mathbf{R}) \xrightarrow{f_2} \mathcal{A}$$

demuestra que $f_2 \circ f_1$ es un *isomorfismo* del grupo (multiplicativo) U sobre el grupo (aditivo) $\mathcal{A}$.

Así 0 es el *ángulo nulo* asociado a $u = 1$,

$$\hat{\theta}_1 = f_2(r_1) = (\widehat{\vec{a}, \vec{b}})$$

es el *ángulo recto* (a, b forman una base ortonormal directa), está asociado a $u = i$.

$$\hat{\theta}_2 = 2\hat{\theta}_1 = f_2(r_2) = (\widehat{\vec{a}, -\vec{a}})$$

es el *ángulo llano,* está asociado a -1. Finalmente

$$r_4 = (r_2)^2 = \mathrm{id}_E \Rightarrow 4\hat{\theta}_1 = 2\hat{\theta}_2 = 0$$

El isomorfismo $u = \alpha + i\beta \rightarrow \hat{\theta}$ permite definir las funciones trigonométricas

$$\hat{\theta} \rightarrow \alpha = \cos \hat{\theta}, \qquad \hat{\theta} \rightarrow \beta = \operatorname{sen} \hat{\theta}.$$

Se ve inmediatamente que

$$(\cos \hat{\theta})^2 + (\operatorname{sen} \hat{\theta})^2 = 1$$
$$\cos(\hat{\theta} + \hat{\theta}') = \cos \hat{\theta} \cos \hat{\theta}' - \operatorname{sen} \hat{\theta} \operatorname{sen} \hat{\theta}'$$
$$\operatorname{sen}(\hat{\theta} + \hat{\theta}') = \operatorname{sen} \hat{\theta} \cos \hat{\theta}' + \operatorname{sen} \hat{\theta}' \cos \hat{\theta}$$

gracias al isomorfismo $f_2 \circ f_1$ y a las propiedades del grupo U se podrá encontrar de nuevo todas las fórmulas de la trigonometría. Pero observemos que las funciones "*coseno*" y "*seno*" que acabamos de definir son unas aplicaciones de $\mathcal{A}$ en $\mathbf{R}$ (de un modo más preciso sobre $[-1, +1]$) y no aplicaciones de $\mathbf{R}$ en $\mathbf{R}$, vamos a definir las funciones trigonométricas que el lector conoce (funciones reales de variable real) gracias a lo que se llama la medida de los ángulos.

OBSERVACION

Si se cambia la orientación de E_2, un ángulo se cambia por su opuesto: sólo el ángulo nulo y el llano son independientes de la orientación escogida en E_2.

c) Medida de los ángulos

Medir una magnitud X es asociarle de manera biunívoca un número real x donde la aplicación $X \to x$ posee ciertas propiedades; en particular si el conjunto $\mathcal{X}$ de magnitudes X está provisto de una adición $(X, Y) \to X + Y$, y posee un elemento neutro O para esta adición, $x + y$ deberá ser la medida de $X + Y$ y 0 la medida O. Si esto fuera posible para los ángulos, designando por θ la medida del ángulo $\widehat{\theta}$, se tendrá en particular $4\widehat{\theta}_1 = 0$ lo que implicaría $\theta_1 = \theta$: el ángulo recto no nulo $\widehat{\theta}_1$, ¡tendría una medida nula!

Procedamos en el otro sentido: supongamos que exista un *homomorfismo* f_0 *del grupo aditivo* **R** *sobre* el grupo multiplicativo U, el diagrama

$$\mathbf{R} \xrightarrow{f_0} U \xrightarrow{f_1} O^+(2, \mathbf{R}) \xrightarrow{f_2} \mathcal{A}$$

demuestra que $f = f_2 \circ f_1 \circ f_0$ será un homomorfismo suprayectivo de **R** sobre $\mathcal{A}$.

Ahora bien, un tal homomorfismo f_0 *existe "admitiremos"* los siguientes resultados:

La aplicación **R** *en* U *definida por*

$$\theta \to e^{i\theta} = \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!}$$

es suprayectiva.

El conjunto de las raíces estrictamente positivas de la ecuación

$$e^{i\theta} = 1$$

tiene un elemento mínimo, se le representa por 2π.

Se demostrará, en efecto, en Análisis que la serie entera de término general $z^n/n!$ es convergente en todo el plano y que si e^z designa la suma

$$e^z e^{z'} = e^{z+z'}$$

se tiene, por tanto,

$$\theta \in R \Rightarrow e^{i\theta}\bar{e}^{i\theta} = e^{i\theta} \cdot e^{-i\theta} = 1 \Rightarrow |e^{i\theta}| = 1$$

en consecuencia f_0 es un homomorfismo del grupo aditivo **R** en el grupo **U**. Admitiremos que es *suproyectivo. El núcleo de* f_0, *o sea,* $f^{-1}(1)$ *es un subgrupo* de **R**, hemos admitido que existía un elemento mínimo estrictamente positivo 2π según los resultados del ejercicio 120 (fin del capítulo 5)

$$f_0^{-1}(1)\, 2\pi\mathbf{Z}.$$

Resulta de lo que hemos demostrado y admitido que $f = f_2 \circ f_1 \circ f_0$ es un *homomorfismo suprayectivo* del *grupo aditivo* **R** *sobre el grupo aditivo de los ángulos* $\mathcal{A}$ y que su núcleo es

$$f^{-1}(0) = 2\pi\mathbf{Z};$$

resulta finalmente que

$$\varphi : \mathbf{R}/2\pi\mathbf{Z} \rightarrow \mathcal{A}$$

es un isomorfismo del *grupo aditivo de los números reales módulo* 2π (V. § 75, ej. 2) *sobre el grupo* $\mathcal{A}$.

Por *definición* $\varphi^{-1}(\hat{\theta}) = \dot{\theta}$ es la *medida* del ángulo $\hat{\theta}$; por abuso de lenguaje diremos que un número real cualquiera θ representante de una clase $\dot{\theta}$ de números reales módulo 2π es *una medida* del ángulo $\hat{\theta}$, si θ_0 es una de ella se tiene

$$\dot{\theta} \equiv \dot{\theta}_0 \pmod{2\pi} \Leftrightarrow \theta = \theta_0 + 2k\pi \qquad (k \in \mathbf{Z}).$$

Así $\pi/2$ es una medída del ángulo recto $\hat{\theta}_1$ y π una medida del ángulo llano $\hat{\theta}_2$.

La medida perteneciente a $[0, 2\pi[$ es llamada algunas veces *medida principal de* $\hat{\theta}$.

Si representamos por θ una medida del ángulo $\hat{\theta}$ asociada a $u = \alpha + i\beta$ ($|u| = 1$) definiremos las aplicaciones de $\mathbf{R}$ en $\mathbf{R}$ (o mejor sobre $[-1, +1]$) por

$$\theta \rightarrow \alpha = \cos\,\theta, \qquad \theta \rightarrow \beta \operatorname{sen}\,\theta$$

resulta de su definición que estas dos funciones son 2π — *periódicas*. Su definición y el homomorfismo f permiten demostrar todas las fórmulas de trigonometría.

EJERCICIO

1. Demostrar que a cuando es un número real no nulo, f_a definido por $f_a(x) = f\left(\frac{2\pi x}{a}\right)$ es un homomorfismo suprayectivo de $\mathbf{R}$ sobre $\mathcal{A}$, todo número real x tal que $f_a(x) = \hat{\theta}$ es una medida de $\hat{\theta}$ respecto a la *base a*. Se puede definir las funciones $\cos_a$ y sen_a correspondientes. Demostrar que

$$\cos_a x + i \operatorname{sen}_a x = e^{\frac{2i\pi x}{a}}.$$

El ángulo de medida 1 respecto a la base a se llama el *ángulo unidad* relativo a la base a. Los ángulos unidades relativos a las bases 2π, 360, 400 se llaman respectivamente el *radián*, el *grado sexagesimal* y el *grado centesimal*.

d) Angulos de dos vectores no nulos o de dos semirrectas de E_2

Dados dos vectores no nulos $\vec{x}$ y $\vec{x}'$ de un espacio vectorial E sobre $\mathbf{R}$ la relación

$$(\exists \lambda \in \mathbf{R}_+^*) \qquad \vec{x}' = \lambda\vec{x}$$

es evidentemente una relación de equivalencia. Se llama *semirrecta abierta* (pasando por 0) de E toda clase de equivalencia en esta relación.

Si E es euclídeo, toda semirrecta abierta (pasando por 0) está asociada de manera biyectiva a un vector unitario único $\vec{u}$: Si $\vec{x}$ es un elemento de la semirrecta abierta D, se tiene

$$\vec{u} = \frac{\vec{x}}{\|\vec{x}\|}.$$

Por "definición" llamaremos ángulo $(\widehat{D_1, D_2})$ *de dos semirrectas abiertas o ángulo* $\left(\widehat{\vec{u}_1, \vec{u}}\right)$ *de un vector* $\vec{x}_1$ *de* D_1 *y de un vector* $\vec{x}_2$ *de* D_2 *el ángulo* $\left(\widehat{\vec{x}_1, \vec{x}_2}\right)$ *de los dos vectores unitarios respectivamente asociados a* D_1 *y a* D_2.

Las medidas de $(\widehat{D_1, D_2})$ o $\left(\widehat{\vec{x}_1, \vec{x}_2}\right)$ serán las de $\left(\widehat{\vec{u}_1, \vec{u}_2}\right)$.

e) **Angulo de dos vectores no nulos de** E_3

Si $\vec{x}_1$ y $\vec{x}_2$ son independientes engendran un plano F pasando por 0 (V. § 137, *b*, observación) la estructura del espacio euclídeo de E_3 induce sobre F una estructura de espacio *euclídeo de dimensión 2, pero la orientación de* F *es independiente de la de* E_3.

Aunque si $\left(\widehat{\vec{x}_1, \vec{x}_2}\right)$ tiene por valor θ para una orientación de F este ángulo tendrá por valor $-\theta$ para la otra orientación, si θ es una de las medidas de este ángulo *sólo* $\cos\theta$ *está definido.*

Si $\vec{x}_1$ y $\vec{x}_2$ no nulos son colineales, es decir, si $\vec{x}_2 = \lambda\vec{x}_1$, $\left(\widehat{\vec{x}_1, \vec{x}_2}\right)$ es el *ángulo nulo* si $\lambda > 0$ y es el *ángulo llano* si $\lambda < 0$: para todo plano F conteniendo $\vec{x}_1$ y $\vec{x}_2$ estos son los dos únicos ángulos independientes de la orientación de F; sus medidas son, respectivamente, 0 y π.

Volvamos al caso general, existe al menos una base ortonormal directa $(\vec{a}, \vec{b}, \vec{c})$ de E_3 tal que

$$\vec{x}_1 = \|\vec{x}_1\|\,\vec{a}$$

y tal que $(\vec{a}, \vec{b})$ sea una base ortonormal de F, existe un par de escalares únicos tales que

$$\vec{x}_2 = \|\vec{x}_2\|\,(\alpha\vec{a} + \beta\vec{b})$$

si θ es una medida del ángulo $\left(\widehat{\vec{x}_1, \vec{x}_2}\right)$ se tendrá

$\alpha = \cos\theta \qquad \beta = \operatorname{sen}\theta \qquad$ si $(\vec{a}, \vec{b})$ es una base directa de F
$\alpha = \cos\theta \qquad \beta = -\operatorname{sen}\theta \qquad$ " " " inversa "

En los dos casos

$$(1) \qquad \vec{x}_1 \cdot \vec{x}_2 = \|\vec{x}_1\|\,\|\vec{x}_2\| \cos\theta$$

observemos que esta fórmula es válida si $\vec{x}_2 = \lambda\vec{x}_1$.

Por otro lado, si ε es igual a + 1 o — 1 según que la base $(\vec{a}, \vec{b})$ de F sea directa o inversa para la orientación escogida en F

$$(2) \qquad \vec{x}_1 \wedge \vec{x}_2 = \|\vec{x}_1\|\,\|\vec{x}_2\|\, \varepsilon \operatorname{sen}\theta\, \vec{c}$$

fórmula también válida cuando $\vec{x}_2 = \lambda\vec{x}_1$, pues en este caso $\vec{x}_1 \wedge \vec{x}_2 = 0$ y $\operatorname{sen}\theta = 0$.

EJERCICIOS

2. Si E_3 esta orientado, demostrar que el escoger una orientación en un plano F (pasando por 0) de E_3 es equivalente a la elección de un vector unitario $\vec{n}$ en $F^{\perp}$.

3. Demostrar que

$$(\vec{x}_1 \cdot \vec{x}_2)^2 + \|\vec{x}_1 \wedge \vec{x}_2\|^2 = \|\vec{x}_1\|^2 \|\vec{x}_2\|^2$$

introduciendo las coordenadas de dos vectores en una base ortonormal de E_3 se obtiene una relación llamada identidad de LAGRANGE.

4. Sea r una rotación de E_3, demostrar que sus valores propios son (con φ no congruente con 0 mod. π)

$$1,\ 1,\ 1 \qquad \text{o} \qquad 1,\ -1,\ -1 \qquad \text{o} \qquad 1,\ e^{i\varphi},\ e^{-i\varphi}.$$

Deducir de lo anterior que existe siempre un vector unitario $\vec{c}$ invariante por r.

Interpretar geométricamente el segundo caso.

En el último caso demostrar que el subespacio ortogonal a $\vec{c}$ es estable por r y que, si se le expresa en una base ortonormal $(\vec{a}, \vec{b})$ tal que $(\vec{a}, \vec{b}, \vec{c})$ sea directa, en esta base

$$M(r) = \begin{pmatrix} \cos\varphi & -\operatorname{sen}\varphi & 0 \\ \operatorname{sen}\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} = A.$$

Si el polinomio característico de r es

$$p_r(X) = p_A(X) = (X - 1)(X^2 - 2\mu X + 1)$$

demostrar que $\cos\varphi = \mu$.

Si $\vec{v}$ es un vector no nulo colineal con $\vec{c}$ y $r(\vec{v}) = \vec{v}'$, demostrar que $\operatorname{sen}\varphi$ tiene el signo del producto mixto $(\vec{v}, \vec{v}', \vec{c})$: φ así determinado es la medida del ángulo de la «rotación r alrededor de $\vec{c}$» (V. § 232 ej. 3).

V. Formas hermitianas. Espacio hermitiano de dimensión n

236. Definiciones diversas

a) Introducción

Hemos visto que sobre $E = \mathbf{C}^n$ espacio vectorial de dimensión n sobre $\mathbf{C}$ la forma bilineal simétrica definida por

$$(x,\ y) \to \sum_{i=1}^{n} x^i y^i$$

es no degenerada sobre E pero que hay en E vectores isótropos no nulos (§ 225, ej. 3); dicho de otra manera $\sum_{i=1}^{n} (x^i)^2 = 0$ se puede verificar para $x \neq 0$.

Consideremos ahora la aplicación f de E^2 en $\mathbf{C}$ definida por

$$(x,\ y) \to f(x,\ y) = \sum_{i=1}^{n} x^i \bar{y}^i$$

tendremos

$$f(x,\ x) = \sum_{i=1}^{n} x^i \bar{x}^i = \sum_{i=0}^{n} |\, x^i \,|^2 \geq 0$$

y

$$f(x,\ x) = 0 \Rightarrow x = 0$$

pero f no es ya una aplicación bilineal de $E \times E$ en $\mathbf{C}$, tenemos en efecto cualquiera que sean x, x', y, y' de E y λ de $\mathbf{C}$

(1) $\quad f(x + x',\ y) = f(x,\ y) + f(x',\ y), \qquad f(\lambda x,\ y) = \lambda f(x,\ y)$

(1′) $\quad f(x,\ y + y') = f(x,\ y) + f(x,\ y'), \qquad f(x,\ \lambda y) = \bar{\lambda} f(x,\ y).$

Este estudio nos induce las definiciones siguientes:

b) Aplicaciones y formas semilineales

Si E *y* F *son dos espacios vectoriales sobre* $\mathbf{C}$ *se llama "aplicación semilineal" de* E *en* F *toda aplicación* u *de* E *en* F *verificando para todo* x *y todo* x' *de* E *y todo* λ *de* $\mathbf{C}$

$$u(x + x') = u(x) + u(x'), \qquad u(\lambda x) = \bar{\lambda} u(x).$$

Si $F = \mathbf{C}$ *se dice que* u *es una "forma semilineal" definida sobre* E.

Sea E un espacio vectorial sobre $\mathbf{C}$, podemos considerar el *conjunto soporte* del espacio vectorial E (V. § 68, *b*), le designaremos con el término

"conjunto E" para simplificar la escritura: podemos proporcionar a este *conjunto* E la *misma suma* que en el *espacio vectorial* E y de la *operación externa siguiente* (con operadores en **C**, V. § 65).

$$(\lambda,\ x) \to \bar{\lambda}x.$$

Se verificará fácilmente que el *conjunto* E provisto de esta suma y de esta multiplicación externa tiene una estructura de *espacio vectorial sobre* **C**, lo designaremos por $\overline{E}$; los *dos conjuntos* E y $\overline{E}$ son los mismos, los *dos grupos aditivos* E y $\overline{E}$ son igualmente los mismos, pero los *dos espacios vectoriales* E y E son distintos. Está claro que las propiedades del automorfismo de **C** definida por $\lambda \to \bar{\lambda}$ (V. § 115)

$$\overline{(\lambda + \mu)} = \bar{\lambda} + \bar{\mu}, \qquad \overline{(\lambda\mu)} = \bar{\lambda}\bar{\mu}, \qquad \overline{(\bar{\lambda})} = \lambda$$
$$\lambda = 0 \Leftrightarrow \bar{\lambda} = 0$$

implican las propiedades siguientes:

— Una parte A de E es libre (resp. ligada) en el espacio vectorial E si y solamente si es libre (resp. ligada) en el espacio vectorial $\overline{E}$.
— Una parte A de E engendra el espacio vectorial E si y solamente si engendra el espacio vectorial E.
— Toda base del espacio vectorial E es una base del espacio vectorial $\overline{E}$ y recíprocamente. Los espacios vectoriales E y $\overline{E}$ son simultáneamente de dimensión finita y su dimensión es la misma.
— A es un subespacio vectorial de E si y solamente si $\overline{A}$ es un subespacio vectorial de $\overline{E}$.
— Finalmente se verifica que toda *aplicación semilineal u de* E *en el espacio vectorial* F *es una aplicación lineal de* E *en el espacio vectorial* $\overline{F}$.

Deducimos de estos resultados que las aplicaciones

$$u : E \to F, \qquad u : E \to \overline{F}$$

tienen las mismas propiedades relativamente a la inyectividad y la suprayectividad. El núcleo de la aplicación semilineal u (de E en $\overline{F}$) es por definición el núcleo de la aplicación lineal u (de E en $\overline{F}$). La imagen u(E) de la aplicación semilineal u (de E en F) es la imagen de la aplicación lineal u (de E en $\overline{F}$); si u(E) es de dimensión finita r diremos que r es el rango de la aplicación semilineal u (de E en F).

El dual E* del espacio vectorial E es el espacio vectorial de las formas lineales definidas sobre E, el dual $\overline{E^*}$ del espacio vectorial $\overline{E}$ es el espacio vectorial de las formas lineales definidas sobre E, es también el *espacio vectorial de las formas semilineales definidas sobre el espacio vectorial* E. Si E es de dimensión finita se tendrá

$$\dim E = \dim E^* = \dim \overline{E^*}.$$

OBSERVACION

Está claro que todo lo que acabamos de decir sobre E, espacio vectorial sobre **C**, puede aplicarse a todo espacio vectorial E sobre el cuerpo conmutativo K, este último estando provisto de un automorfismo *involutivo* de K (§ 16, ej. 6), es decir, teniendo las mismas propiedades que la aplicación $\lambda \to \bar{\lambda}$ de **C** en **C**.

c) Formas sesquilineales

Si E *y* F *son dos espacios vectoriales sobre* **C** *se dice que una aplicación f de* E × F *en* **C** *es una "forma sesquilineal" si verifica las relaciones* (1) *y* (1'): cualesquiera que sean x, x' de E, y e y' de F y λ de **C**

$$(1) \qquad f(x + x', y) = f(x, y) + f(x', y), \qquad f(\lambda x, y) = \lambda f(x, y)$$

$$(1') \qquad f(x, y + y') = f(x, y) + f(x, y'), \qquad f(x, \lambda y) = \bar{\lambda} f(x, y).$$

Dicho de otra forma f es sesquilineal si y solamente si

$f_x = f(x, \cdot)$ es una forma *semilineal* sobre F
$f_y = f(\cdot, y)$ es una forma *lineal* sobre E

donde

$$f_x \in \mathrm{F}^*, \qquad f_y \in \mathrm{E}^*.$$

Si E y F son de dimensión finita, referidos respectivamente a las bases (a_i) $(1 \leq i \leq m)$ y (b_j) $(1 \leq j \leq n)$ poniendo

$$x = \sum_{i=1}^{m} x^i a_i, \qquad y = \sum_{j=1}^{n} y^j b_j, \qquad f(a_i, b_j) = \alpha_{ij} \in \mathbf{C}$$

obtendremos como en el párrafo 221

$$f(x, y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} x^i \bar{y}^j.$$

La matriz $\mathrm{A} = (\alpha_{ij})$ matriz compleja de m columnas y n filas se le llama *asociada a la forma sesquilineal f;* con las notaciones habituales $\mathrm{X} = \mathrm{M}(x, (a_i))$, $\mathrm{Y} = \mathrm{M}(y, (b_j))$ tendremos

$$f(x, y) = {}^t\overline{\mathrm{Y}}\mathrm{A}\mathrm{X} = {}^t\mathrm{X}\,{}^t\mathrm{A}\overline{\mathrm{Y}}.$$

EJERCICIOS

1. Demostrar que el conjunto de las formas sesquilineales sobre E × F es un espacio vectorial sobre **C**. ¿Cuál es su dimensión si dim $\mathrm{E} = m$, dim $\mathrm{F} = n$?
2. Demostrar con las notaciones del § 221 que

$$\mathrm{A}' = {}^t\overline{\mathrm{Q}}\mathrm{A}\mathrm{P}.$$

d) Formas hermitianas

Supongamos E = F, se puede entonces considerar simultáneamente $f(x, y)$ y $f(y, x)$; se dice que la *forma sesquilineal f definida sobre* E × E *es "hermitiana"* (V. observación siguiente) *si y solamente si*

$$(\forall (x, y) \in E^2) \qquad f(y, x) = \overline{f(x, y)}.$$

A esta relación se le llama algunas veces "*simetria hermitiana*" que es un abuso de lenguaje, ya que esta relación no es una "simetria", x e y tienen diferentes papeles.

Se dirá que f es una "*forma hermitiana sobre* E". Resulta inmediatamente de esta definición que para todo x, $f(x, x)$ es real. Se llamará *forma cuadrática hermitiana* asociada a la forma hermitiana f la aplicación q de E en **R** definida por

$$(2) \qquad (\forall x \in E) \qquad q(x) = f(x, x).$$

Utilizando esta definición y las propiedades de la forma hermitiana f se verifica inmediatamente que

$$(2') \quad (\forall (x, y) \in E^2) \quad f(x, y) = \frac{q(x+y) - q(x-y)}{4} + i\,\frac{q(x+iy) - q(x-iy)}{4}$$

luego si para todo x, $q(x) = f(x, x) = g(x, x)$, con f y g dos formas hermitianas sobre E se tiene $f = g$; resulta de lo anterior igualmente que $f = 0$ es equivalente a $q = 0$.

Sea (a_i) $(1 \leq i \leq n)$ una base de E supuesta de dimensión n sobre **C**, la matriz $A = (\alpha_{ij})$ asociada a la forma hermitiana f relativamente a la base (a_i) es tal que para todo par (i, j)

$$\alpha_{ji} = f(a_j, a_i) = \overline{f(a_i, a_j)} = \overline{\alpha_{ij}}$$

dicho de otra manera

$${}^tA = \overline{A}$$

hemos dicho ya (V. § 154, *c*) que tal matriz cuadrada con elementos complejos se llamaba *matriz hermitiana*. Se vé inmediatamente que f es hermitiana si y solamente si la matriz asociada en una base cualquiera de E es hermitiana. Con las mismas notaciones se tiene

$$(3) \qquad f(x, y) = {}^t\overline{Y}AX$$

$$(3') \qquad q(x) = f(x, x) = {}^t\overline{X}AX.$$

OBSERVACION SOBRE LA TERMINOLOGIA(*)

La terminología no está aún completamente fijada. Hemos utilizado «aplicación (o forma) semilineal», «forma sesquilineal», «forma hermitiana», los términos que parecen se imponen añadiéndoles el término de «*forma cuadrática hermitiana*». Tenemos así el cuadro de correspondencias siguiente que se comprende fácilmente:

(*) *N. del T.* — Estas consideraciones las hacemos extensivas al idioma español.

Conjunto de definición		
E	aplicación (forma) *lineal*	aplicación (forma) *semilineal*
E × E	forma *bilineal*	forma *sesquilineal*
E × E	forma *bilineal simétrica*	forma *hermitiana*
E	forma *cuadrática*	forma *cuadrática hermitiana*

Naturalmente para las aplicaciones de la última columna $K = \mathbf{C}$. Ciertos autores llaman forma hermitiana a lo que nosotros hemos llamado forma cuadrática hermitiana. Los físicos dicen en general «hermítica» en lugar de «hermitiana».

EJEMPLOS Y EJERCICIOS

3. La forma que ha servido de introducción a este párrafo

$$E = \mathbf{C}^n \qquad f(x, y) = \sum_{i=1}^{n} x^i \bar{y}^i$$

es hermitiana, su matriz asociada en la base canónica de $\mathbf{C}^n$ es $\mathbf{I}_n$.

4. Siendo E el espacio vectorial de las funciones complejas de variable real continuas sobre $[\alpha, \beta]$, la aplicación de E × E en **C** definida por

$$f(x, y) = \int_\alpha^\beta x(t)\overline{y(t)}\, dt$$

es una forma hermitiana sobre E.

5. Demostrar con la ayuda del ejercicio 2 anterior, que si la matriz asociada a una forma sesquilineal definida sobre E × E es hermitiana en una base, lo es en toda base de E.

237. Estudio de las formas hermitianas sobre E, espacio vectorial de dimensión finita sobre C

a) Introducción

Todo lo que hemos dicho en los §§ 224 hasta el 231 sobre las formas bilineales simétricas se aplica con o sin modificación a las formas hermitianas; no tenemos intención de entretener al lector con el detalle de todas las demostraciones: daremos íntegramente sólo aquellas que están bastante modificadas por el paso de la "simetría" (sin más) a la "simetría hermitiana", contentándonos en dar algunas indicaciones para las que no son modificadas en su principio. Sin embargo, aconsejamos vivamente al lector el redactar completamente estas demostraciones como ejercicio, para ver mejor los cambios efectuados en el paso del caso "bilineal simétrico" al caso "hermitiano".

Hagamos primero algunas observaciones. $\lambda = 0$ siendo equivalente a $\bar{\lambda} = 0$ tenemos

$$f(x, y) = 0 \Leftrightarrow f(y, x) = \overline{f(x, y)} = 0$$

si una de estas relaciones se verifica diremos que x e y de E son *ortogonales* relativamente a f; de ello se deduce inmediatamente la definición de *subespacios ortogonales* y de *ortogonal* $F^{\perp}$ de un subespacio F de E (relativamente a f). De lo anterior se obtiene igualmente la definición de los elementos x de E, isótropos (relativamente a f) por $f(x, x) = 0$, así como la definición de un *subespacio isótropo* F de E: existe $x \neq 0$ isótropo ortogonal a f todo entero.

Una base (a_i) de E será *ortogonal* relativamente a f si

$$i \neq j \Rightarrow f(a_i, a_j) = 0$$

y *ortonormal* relativamente a f si es ortogonal y si

$$(i = 1, \ldots, n) \qquad f(a_i, a_i) = 1.$$

Por otro lado la forma $f_x = f(x, \cdot)$ es semilineal, por el contrario la forma $f_y = f(\cdot, y)$ es lineal pues f_y pertenece a E^*, designaremos aquí por φ *únicamente la aplicación de* E *en* E^* definida por

$$(\forall y \in E) \qquad \varphi(y) = f_y = f(\cdot, y)$$

vamos a ver que esta aplicación es semilineal.

Recordemos en fin que en el caso hermitiano $f(x, x)$ es *real* cualquiera que sea x de E.

b) Núcleo de una forma hermitiana. Forma hermitiana no degenerada

Se demostrará como en el § 222 que, cualesquiera que sean y_1 e y_2 de E

$$\varphi(y_1 + y_2) = f_{y_1+y_2} = \varphi(y_1) + \varphi(y_2) = f_{y_1} + f_{y_2}$$

por otra parte, cualesquiera que sean x e y de E y λ de **C**

$$f_{\lambda y}(x) = f(x, \lambda y) = \bar{\lambda} f(x, y) = \bar{\lambda} f_y(x)$$

es decir,

$$f_{\lambda y} = \bar{\lambda} f_y \Leftrightarrow \varphi(\lambda y) = \bar{\lambda}\varphi(y)$$

en consecuencia φ es una *aplicación semilineal* de E en E^*, es pues una *aplicación* lineal de E en $\overline{E^*}$ (V. § 236, *b*).

El conjunto de los y tales que $\varphi(y) = 0$, es decir, el núcleo de la aplicación lineal φ se le llamará también el *núcleo de la forma hermitiana* f.

Si $\operatorname{Ker} \varphi = \operatorname{Ker} f = \{0\}$, $\varphi(y) = 0$ implica $y = 0$, es decir,

$$(1) \qquad [(\forall x \in E) \qquad [\varphi(y)](x) = f_y(x) = f(x, y) = 0] \Rightarrow y = 0.$$

La relación $f(y, x) = \overline{f(x, y)}$ demuestra que la implicación (1) es equivalente a la implicación (2)

$$(2) \qquad [(\forall y \in E) \qquad f(x, y) = 0] \Rightarrow x = 0.$$

Seguidamente si $\operatorname{Ker} \varphi = \{0\}$, φ aplicación lineal de E en $\overline{E^*}$ es inyectiva, luego biyectiva puesto que $\dim \overline{E^*} = \dim E = n$, y recíprocamente, φ biyectiva implica (1). Se dice en este caso que f es *no degenerada* sobre E.

En fin, sea $A = (\alpha_{ij})$ la matriz asociada a la forma hermitiana f, tendremos

$$f(x, y) = 0 \Leftrightarrow \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij}x^i\bar{y}^j = 0$$

la relación (I) es equivalente a la siguiente: el sistema lineal cuyas incógnitas son $\bar{y}^1, \ldots, \bar{y}^n$

$$(i = 1, \ldots, n) \qquad \sum_{j=1}^{n} \alpha_{ij}\bar{y}^j = 0$$

sólo tiene la solución trivial ($\bar{y}^1 = \ldots = \bar{y}^n = 0$) para ello es necesario y suficiente que det $(\alpha_{ij}) \neq 0$, de donde:

TEOREMA 6'. — *Siendo f una forma hermitiana sobre* E, *espacio vectorial de dimensión n sobre* **C** *y (α_{ij}) la matriz asociada a f en una base de* E, *las propiedades siguientes son equivalentes*

1.ª *f no es degenerada sobre* E.
2.ª *$f(x, y) = 0$ para todo x de* E *implica $y = 0$.*
2.ª *$f(x, y) = 0$ para todo y de* E *implica $x = 0$.*
4.ª *La aplicación φ de* E *en* E* *definida por $\varphi(y) = f_y = f(\cdot , y)$ es biyectiva.*
5.ª det $(\alpha_{ij}) \neq 0$.

En particular si f es no degenerada, la aplicación semilineal φ de E en E* es biyectiva, cualquiera que sea l^* de E*, existe y único de E, tal que $l^* = \varphi(y) = f_y$, es decir, existe y único de E, tal que

$$(\forall\, x \in E) \qquad l^*(x) = f_y(x) = f(x, y).$$

La dimensión r de $\varphi(E)$, es decir, el *rango* de φ aplicación lineal de E en E* se llama también el *rango de la forma hermitiana* f. El *núcleo* de f, es decir, el núcleo de φ es de dimensión $n - r$.

Si $r < n$, es decir, si Ker $\varphi \neq \{0\}$ se dice que la forma hermitiana f está *degenerada* sobre E.

Volvamos al caso en que f *es no degenerada sobre* E, la biyección $y \to \varphi(y)$ nos permite identificar E y E* poniendo $y = \varphi(y)$ se tendrá, pues, para todo x y para todo y de E

$$\langle x, y \rangle = \langle x, \varphi(y) \rangle = [\varphi(y)]\,(x) = f_y(x) = f(x, y).$$

Esta identificación nos permite como en el § 225, *c*, confundir:

— La ortogonalidad entre x de E e y de E*: $\langle x, y \rangle = 0$.
— La ortogonalidad en E relativamente a f: $f(x, y) = 0$.

Se deduce inmediatamente:

TEOREMA 8'. — *Si f es una forma hermitiana no degenerada sobre* E, *espacio vectorial de dimensión finita y* $F^\perp$ *el ortogonal de* F *relativamente a f, para subespacio* F *de* E *se tiene*

$$\dim F^\perp = \dim E - \dim F, \qquad (F^\perp)^\perp = F.$$

EJERCICIO

Demostrar los resultados siguientes:

TEOREMA 7′: Siendo E un espacio vectorial de dimensión finita sobre **C**, para toda forma hermitiana f sobre E y todo subespacio vectorial F de E las propiedades siguientes son equivalentes:

1) La restricción de f a F es no degenerada.
2) $F \cap F^{\perp} = \{0\}$.
3) F no es isótropo.
4) $E = F \oplus F^{\perp}$.

c) Existencia de bases ortogonales

La demostración de la existencia de bases ortogonales (relativamente a f) para toda forma bilineal simétrica es basada en el § 226, *b*, en el hecho que si a_1 no es isótropo $F = Ka_1$ es tal que $E = F \oplus F^{\perp}$. La demostración será exactamente la misma para f hermitiana tomando $F = \mathbf{C}a_1$, el resultado indicado es una consecuencia del teorema 7′ (ver ejercicio anterior). Para el lector que no hubiera demostrado el teorema 7′ damos una demostración independiente de este teorema:

Sea a_1, tal que $f(a_1, a_1) \neq 0$, tomemos $F = \mathbf{C}a_1$ y completemos a_1 con $a_2, \ldots, a_n$ para obtener una base (a_i) de E, en esta base tendremos

$$f(x, y) = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij}x^i\bar{y}^j \qquad f(a_1, a_1) = \alpha_{11} \neq 0$$

el ortogonal de $y = a_1$ ($y^1 = 1$, $y^2 = \ldots = y^n = 0$), es decir, $F^{\perp}$ tiene por ecuación cartesiana (ver § 151, ej. 1)

$$\alpha_{11}x^1 + \ldots + \alpha_{n1}x^n = 0$$

donde no todos los coeficientes α_{i1} son nulos ($\alpha_{11} \neq 0$), $F^{\perp}$ es, pues, un ***hiperplano*** de E. Por otra parte $F^{\perp}$ no contiene a_1, pues por $x^1 = 1$, $x^2 = \ldots = x^n = 0$ se tendría, si $F^{\perp}$ contuviera a_1, $\alpha_{11} = 0$ lo que es contrario a la hipótesis, luego

$$\dim F = 1, \qquad \dim F^{\perp} = n - 1, \qquad F \cap F^{\perp} = \{0\}$$

es decir,

$$E = F \oplus F^{\perp}$$

la demostración se acaba como en el § 226, *b*, de donde:

TEOREMA 9′. — *Para todo espacio vectorial* E *de dimensión* n *sobre* **C** *existen bases ortogonales relativamente a toda forma hermitiana* f *sobre* E. *De un modo más preciso siendo* f *una forma hermitiana de rango* r, *existen bases de* E *tales que*

$$i \neq j \Rightarrow f(a_i, a_j) = 0$$
$$1 \leq i \leq r \Rightarrow f(a_i, a_i) \neq 0$$
$$r + 1 \leq i \leq n \Rightarrow f(a_i, a_i) = 0.$$

En una tal base

$$f(x, y) = \sum_{i=1}^{r} \alpha_{ii} x^i \bar{y}^i, \qquad q(x) = \sum_{i=1}^{r} \alpha_{ii} x^i \bar{x}^i.$$

Las $\alpha_{ii} (1 \leq i \leq r)$ al ser números reales no nulos, como en el § 227, se demostrará:

TEOREMA 15'. — *Si q es la forma cuadrática hermitiana asociada a una forma hermitiana f sobre E, espacio vectorial de dimensión n sobre* **C**.

1. *Existe una base (a_i) ortogonal relativamente a f y dos enteros naturales s y t tales que respecto a esta base (a_i) se tiene*

$$q(x) = \sum_{i=1}^{s} x^i \bar{x}^i - \sum_{j=s+1}^{s+t} x^j \bar{x}^j \qquad (s + t = r = rg(f)).$$

2. *Si existe una segunda base (a'_i) y dos enteros naturales s' y t' tales que respecto a (a'_i) se tenga*

$$q(x) = \sum_{i=1}^{s'} x'^i \bar{x}'^i - \sum_{j=s'+1}^{s'+t'} x'^j \bar{x}'^j \qquad (s' + t' = r = rg(f))$$

entonces $s = s'$, $t = t'$ (ley de inercia).

(s, t) es la *signatura* de la forma hermitiana f.

238. Formas hermitianas positivas y no degeneradas positivas. Espacio hermitiano de dimensión n

a) Formas hermitianas positivas. Desigualdad de Schwarz

Recordemos que para toda forma hermitiana f sobre E espacio vectorial sobre **C**, $f(x, x)$ es real podemos, pues, enunciar:

DEFINICIÓN 7'. — *Diremos que una forma hermitiana f y la forma cuadrática hermitiana q asociada son positivas si*

$$(\forall x \in E) \qquad q(x) = f(x, x) \geq 0.$$

Siendo f hermitiano positivo para toda pareja (x, y) de E^2 y todo λ de **C**, tendremos

$$q(x + \lambda y) = f(x + \lambda y, x + \lambda y) = f(x, x) + \bar{\lambda} f(x, y) + \lambda f(y, x) + \lambda\bar{\lambda} f(y, y) \geq 0$$

de donde, poniendo $\alpha = f(x, x)$, $\beta = f(x, y)$, $\gamma = f(y, y)$ y teniendo en cuenta que $\alpha \geq 0$, $\gamma \geq 0$, tendremos

$$\alpha\gamma + \bar{\lambda}\beta\gamma + \lambda\bar{\beta}\gamma + \lambda\bar{\lambda}\gamma^2 \geq 0$$
$$(\lambda\gamma + \beta)(\bar{\lambda}\gamma + \bar{\beta}) + \alpha\gamma - \beta\bar{\beta} = |\lambda\gamma + \beta|^2 + \alpha\gamma - |\beta|^2 \geq 0.$$

Si $\gamma \neq 0$ haciendo $\lambda = -\beta/\gamma$ se obtiene $\alpha\gamma - |\beta|^2 \geq 0$.
Si $\gamma = 0$ no se puede tener para todo λ

$$\alpha + \bar{\lambda}\beta + \lambda\bar{\beta} = \alpha + 2\Re\bar{\lambda}\beta \geq 0$$

más que si β es nulo, el segundo término puede ser un número real cualquiera si $\beta \neq 0$, luego, en particular, estrictamente inferior a $-\alpha$; se obtiene también así $\alpha\gamma - |\beta|^2 \geq 0$, de donde:

TEOREMA 14'. — *Si f es una forma hermitiana positiva sobre* E, *espacio vectorial sobre* **C**, *se tiene*

$$(\forall (x, y) \in E^2) \qquad |f(x, y)|^2 \leq f(x, x)f(y, y).$$

(Desigualdad de Schwarz).

Se deduce inmediatamente

$$\begin{aligned} q(x + y) &= f(x + y, x + y) = f(x, x) + f(x, y) + \overline{f(x, y)} + f(y, y) \\ &= f(x, x) + 2\Re f(x, y) + f(y, y) \leq f(x, x) + 2|f(x, y)| + f(y, y) \\ &\leq q(x) + 2\sqrt{q(x)q(y)} + q(y) = [\sqrt{q(x)} + \sqrt{q(y)}]^2. \end{aligned}$$

COROLARIO 1. — *Siendo q la forma cuadrática hermitiana asociada a la forma hermitiana positiva f*

$$(\forall (x, y) \in E^2) \qquad \sqrt{q(x + y)} \leq \sqrt{q(x)} + \sqrt{q(y)}.$$

Como en el § 230 se deduce igualmente de la desigualdad de SCHWARZ el resultado siguiente:

COROLARIO 2. — *El núcleo de una forma hermitiana f sobre* E, *espacio vectorial sobre* **C** *es el conjunto de los vectores isótropos de* E *relativamente a f.*

Resulta que si la forma hermitiana f positiva es *no degenerada,* su núcleo está reducido a $\{0\}$, luego $x = 0$ es el único vector isótropo. Se dice entonces que la forma hermitiana f (o la forma cuadrática asociada) es *no degenerada positiva* (ciertos autores emplean el término "definida positiva", aquí no lo haremos).

EJEMPLOS Y EJERCICIOS

1. Sea $E = \mathbf{C}^n$, la forma hermitiana definida en la base canónica de E por

$$f(x, y) = \sum_{i=1}^{m} x^i\bar{y}^i \qquad (m \leq n)$$

es positiva; es no degenerada positiva si y solamente si $m = n$.

2. La forma hermitiana definida en el ejercicio 4 del § 236 es no degenerada positiva.

b) Formas hermitianas no degeneradas positivas

Para toda *forma hermitiana positiva sobre* E la aplicación definida por

$$(\forall\, x \in \mathrm{E}) \qquad x \to \sqrt{q(x)} = \sqrt{f(x,\ x)}$$

es una aplicación de E en $\mathbf{R}_+$. Para todo λ de $\mathbf{C}$, se tiene

$$\sqrt{q(\lambda x)} = \sqrt{f(\lambda x,\ \lambda x)} = \sqrt{\lambda\bar{\lambda} f(x,\ x)} = |\lambda|\sqrt{q(x)}$$

y para todo par $(x,\ y)$ de E^2

$$\sqrt{q(x+y)} \leq \sqrt{q(x)} + \sqrt{q(y)}.$$

Si se supone, además, *f no degenerada* el corolario 2 anterior muestra que

$$q(x) = f(x,\ x) = 0 \Rightarrow x = 0$$

de donde:

Teorema 16′. — *Si q es la forma hermitiana asociada a una forma hermitiana no degenerada positiva sobre* E, *espacio vectorial sobre* $\mathbf{C}$, *la aplicación de* E *en* $\mathbf{R}_+$ *definida por*

$$x \to \sqrt{q(x)}$$

es una norma sobre E.

Definición 9′. — E *espacio vectorial sobre* $\mathbf{C}$, *provisto de una norma* $x \to \sqrt{q(x)}$, *si q es una forma cuadrática hermitiana no degenerada positiva sobre* E, *se llama «espacio prehilbertiano complejo» y «espacio hermitiano» si es de dimensión finita.*

Volvamos al teorema 15′ (ver § 237) sobre las bases ortogonales relativamente a la forma hermitiana sobre E, de dimensión n

$$f(x,\ x) = q(x) = \sum_{i=1}^{s} x^i x^i - \sum_{j=s+1}^{s+t} x^j x^j$$

$t = 0$ si y solamente si la forma f es positiva. En este caso $r = \mathrm{rg}\ (f) = s$, luego $s = n$ *si y sólo si* f es no degenerada positiva: en este caso la base hallada es ortonormal, de donde:

Corolario del teorema 15′. — *Siendo* E *un espacio vectorial de dimensión finita sobre* $\mathbf{C}$ *existen bases de* E *ortonormales relativamente a toda forma hermitiana no degenerada positiva.*

c) Espacio hermitiano de dimensión *n*

Para toda forma f, hermitiana no degenerada positiva sobre E, en relación a toda base de E ortonormal relativamente a f, se tiene

$$f(x,\ y) = \sum_{i=1}^{n} x^i \bar{y}^i, \qquad q(x) = f(x,\ x) = \sum_{i=1}^{n} x^i \bar{x}^i = \sum_{i=1}^{n} |\, x^i\,|^2.$$

Esto nos permite decir: *existe una sola estructura de espacio hermitiano de dimensión n.* Para estudiar un tal espacio E escogeremos una forma hermitiana no degenerada positiva f_0 sobre E que se llama *forma fundamental* del espacio hermitiano E; las bases ortonormales de E serán bases ortonormales relativamente a f_0. En lugar de $f_0(x, y)$ se escribe a menudo

$$(x \mid y)$$

y se dice que $(x \mid y)$ es el *producto hermitiano* de x e y.

Llamaremos *norma hermitiana* de x el número real positivo

$$\|x\| = \sqrt{(x|x)}.$$

En una base ortonormal cualquier relativamente a f_0, se tiene

$$(x|y) = x^1\overline{y^1} + \dots + x^n\overline{y^n}, \qquad \|x\| = \sqrt{|x^1|^2 + \dots + |x^n|^2}.$$

OBSERVACION

En lugar de «producto hermitiano» algunos autores emplean los términos «producto escalar» o bien «producto escalar hermitiano». Los físicos emplean «producto hermítico»; observemos igualmente, para que el lector esté prevenido, que en muchos libros de Física se escribe

$$(x \mid y) = \overline{x^1}y^1 + \dots + \overline{x^n}y^n.$$

EJEMPLOS Y EJERCICIOS

3. Demostrar que toda familia (x_i) de vectores no nulos ortogonales dos a dos, de un espacio hermitiano es libre.
4. Aplicar al espacio hermitiano los resultados de los ejercicios 8 y 9 del § 231 (Ortonormalización de SCHMITT).
5. Extender al espacio hermitiano el resultado del ejercicio 11 del § 231 (Teorema de PITÁGORAS).

239. Adjunto de un endomorfismo en un espacio hermitiano. Grupo unitario

Si u y v son dos operadores de un espacio hermitiano E, se demostrará como en el § 227

$$[(\forall (x, y) \in E^2) \quad (u(x) \mid y) = (v(x) \mid y)] \Rightarrow u = v$$
$$[(\forall (x, y) \in E^2) \quad (x \mid u(y)) = (x \mid v(y))] \Rightarrow u = v.$$

a) Adjunto de un endomorfismo

La forma fundamental f_0 es no degenerada, la aplicación asociada φ de E en E* es biyectiva, φ es un isomorfismo de E sobre $\overline{E^*}$, como en el § 227 vemos que a un endomorfismo de E se le puede asociar un endomorfismo único de E, u^* llamado *adjunto* de u; con la notación $f_0(x, y) = (x \mid y)$, se tiene

$$[\forall (x, y) \in E^2] \qquad (u(x) \mid y) = (x \mid u^*(y))$$

es inmediato de verificar que

$$(u+v)^* = u^* + v^*, \qquad (u \circ v)^* = v^* \circ u^*, \qquad (u^*)^* = u, \qquad \text{rg } u^* = \text{rg } u.$$

Se ve igualmente que $(\mathrm{id}_E)^* = \mathrm{id}_E$ y que si u es inversible también lo es u^* y que $(u^*)^{-1} = (u^{-1})^*$. Por el contrario, tendremos para todo par de E^2 y todo λ de $\mathbf{C}$

$$\begin{aligned}(x \mid (\lambda u)^*(y)) &= ((\lambda u)(x) \mid y) = (\lambda u(x) \mid y) = \lambda(u(x) \mid y) \\ &= \lambda(x \mid u^*(y)) = (x \mid \bar{\lambda} u^*(y)) = (x \mid (\bar{\lambda} u^*)(y))\end{aligned}$$

de donde

$$(\lambda u)^* = \bar{\lambda} u^*.$$

Sea por otra parte una base ortonormal (a_i) de E; para todo par de enteros (i, j) de $[1, n]^2$ tenemos $(a_i \mid a_j) = \delta_{ij}$; de donde escribiendo

$$A = M[u, (a_i)] = (\alpha_i^j), \qquad B = M[u^*, (a_i)] = (\beta_i^j)$$

es decir,

$$(i = 1, \ldots, n) \qquad u(a_i) = \sum_{j=2}^{n} \alpha_i^j a_j, \qquad u^*(a_i) = \sum_{j=2}^{n} \beta_i^j a_j$$

se obtiene para todo i y para todo j de $[1, n]$

$$\beta_i^j = (u^*(a_i) \mid a_j) = (a_i \mid u(a_j)) = (u(a_j) \mid a_i) = \alpha_j^i$$

en consecuencia

$$B = {}^t\bar{A}$$

ya hemos dicho en el § 154, *c*, que la matriz ${}^t\bar{A}$ se llamaba la *adjunta* de A ($A \in \mathbf{M}_n(\mathbf{C})$) como hemos dicho en aquel párrafo se la designa $\mathbf{A}^*$; luego si (a_i) es ortonormal

$$M[u^*, (a_i)] = [M[u, (a_i)]]^*$$

se ve, además, que

$$\det u^* = \det A^* = \overline{\det A} = \overline{\det u}$$

de donde:

TEOREMA 11'. — *Si* E *es un espacio hermitiano de dimensión n, a todo endomorfismo u de* E *se puede asociar un endomorfismo único u^* de* E, *llamado «adjunto» de u, tal que*

$$[\forall (x, y) \in E^2] \qquad (u(x) \mid y) = (x \mid u^*(y)).$$

Cualesquiera que sean u y v de $\mathcal{L}(E)$ y λ de $\mathbf{C}$

$$\begin{gathered}(u+v)^* = u^* + v^*, \qquad (\lambda u)^* = \bar{\lambda} u, \qquad (u \circ v)^* = v^* \circ u^* \\ (u^*)^* = u, \qquad \text{rg } u^* = \text{rg } u, \qquad \det u^* = \overline{\det u}.\end{gathered}$$

Si u es inversible, u^* lo es también y $(u^*)^{-1} = (u^{-1})^*$. Las matrices de u y u^* respecto a una misma base ortonormal de E *son adjuntas la una de la otra, es decir,*

$$M(u^*) = {}^t(\overline{M(u)}).$$

b) **Grupo unitario**

Sea u un endomorfismo de E, tal que

(1) $$[\forall (x, y) \in E^2] \qquad (u(x) \mid u(y)) = (x \mid y)$$

se dice que *conserva* el producto hermitiano, es decir, la forma fundamental f_0; vemos de lo anterior que conserva la forma cuadrática hermitiana q_0 asociada a f_0, es decir, (2) implica (2′)

(2) $$(\forall x \in E) \qquad \| u(x) \| = \| x \|$$

recíprocamente la fórmula (2′) del § 236, *d*) demuestra que (2) implica (1).

Como en el § 228, *a*) se demostrarán las propiedades siguientes de un tal endomorfismo, de donde:

TEOREMA 12′. — *Si* E *es un espacio hermitiano de dimensión finita, para todo endomorfismo u de* E *las cinco propiedades siguientes son equivalentes*

1. $[\forall (x, y) \in E^2] \qquad (u(x) \mid u(y)) = (x \mid y).$
2. $(\forall x \in E) \qquad \| u(x) \| = \| x \|.$
3. $u^* \circ u = \mathrm{id}_E.$
4. $u \circ u^* = \mathrm{id}_E.$
5. *u es inversible y* $u^{-1} = u^*$.

Todo endomorfismo del espacio hermitiano E, verificando una de estas cinco propiedades, se llama *automorfismo unitario* de E; se le llama también *operador unitario* (que es un endomorfismo, resulta del hecho de que es unitario). Se dice también que u es un *automorfismo* de la forma hermitiana f_0 o de la forma cuadrática q_0 hermitiana asociada. Sea A la matriz asociada a un operador unitario, respecto a una base ortonormal de E, el teorema 12′ muestra que

$$A^*A = I_n \Leftrightarrow A^{-1} = A^* = {}^t\overline{A}.$$

De una manera general toda matriz de $\mathbf{M}_n(\mathbf{C})$ verificando una de estas condiciones se llama *matriz unitaria* de orden n; se verá exactamente como en el § 228 que *toda matriz de paso de una base ortonormal a otra base ortonormal de un espacio hermitiano es unitaria.*

Se observa también que toda *matriz unitaria real* es *ortogonal.*

c) **Grupo unitario**

Como en el § 228, *b*), tenemos el resultado siguiente:

COROLARIO. — *Los operadores unitarios de un espacio hermitiano de dimensión n describen un subgrupo del grupo* $\mathbf{GL}_n(\mathbf{C})$ *llamado «grupo unitario» y representado* $\mathbf{U}(n, \mathbf{C})$.

Por otra parte, todo operador unitario o toda matriz unitaria

$$\det (u^* \circ u) = \overline{\det u} \cdot \det u = \det (\mathrm{id}_E) = 1$$

luego: *el determinante de todo operador unitario, o de toda matriz unitaria es un número complejo de módulo* 1.

Está claro que el conjunto de los operadores (o de las matrices) unitarios de determinante + 1 es un subgrupo de **U**(n, **C**), se le llama el *grupo unitario especial* y se le representa **SU**(n, **C**).

EJERCICIOS

1. Demostrar que **SU**(n, **C**) es un subgrupo distinguido de **U**(n, **C**) (considerar el homorfismo $u \to \det u$).
2. Demostrar que todo valor propio de un operador unitario es de módulo 1.

240. Operadores hermitianos

Se llama *operador hermitiano* de un espacio hermitiano E todo operador *autoadjunto*, es decir, tal que $u = u^*$, u^* siendo el adjunto de u relativamente a la forma fundamental f_0 de E.

a) Primeras propiedades de los operadores hermitianos

Como en el § 229 se verá que $u = u^*$ es equivalente a

$$(1) \qquad [\forall (x, y) \in E^2] \qquad (u(x) \mid y) = (x \mid u(y))$$

u será hermitiano si y sólo si en una *base ortonormal*

$$M(u) = [M(u)]^* = {}^t\overline{M(u)}$$

es decir (ver § 154, *b* y 236, *d*), si y solamente si $M(u)$ es *hermitiano*.

Se ve igualmente que para todo x de E

$$(x \mid u(x)) = (u(x) \mid x) = \overline{(x \mid u(x))}$$

luego $(x \mid u(x))$ o $(u(x) \mid x)$ es real. Recíprocamente sea u un endomorfismo del espacio hermitiano, tal que $(x \mid u(x))$ sea real para todo x de E, consideremos la aplicación f_u de E^2 en **C** definida por

$$(x, y) \to f_u(x, y) = (x \mid u(y))$$

se ve inmediatamente que f_u es *lineal* en x y *semilineal* en y, es, pues, una *forma sesquilineal*, además, al ser $f_u(x, x)$ real para todo x de E, desarrollando $f_u(x + y, x + y)$ y $f_u(x + iy, x + iy)$ se demostrará que $f_u(x, y) + f_u(y, x)$ es *real* y que $f_u(x, y) - f_u(y, x)$ es *complejo puro*; de ello resulta

$$f_u(y, x) = \overline{f_u(x, y)}$$

luego f_u es *hermitiana*, se tiene, pues,

$$(x \mid u(y)) = \overline{(y \mid u(x))} = (u(x) \mid y)$$

y u es un operador hermitiano, de donde:

TEOREMA 13. — *Si* E *es un espacio hermitiano de dimensión n, para todo operador u de* E *las propiedades siguientes son equivalentes:*

1. $u = u^*$.
2. $[\forall (x, y) \in E^2] \qquad (u(x) \mid y) = (x \mid u(y))$.
3. $(\forall\, x \in E) \qquad (x \mid u(x)) \in \mathbf{R}$.
4. *La matriz asociada a u en una base ortonormal de* E *es hermitiana.*

Observemos finalmente que una *matriz simétrica real* es un caso particular de *matriz hermitiana.*

b) Estudio de una biyección

Designemos por $\mathcal{H}(E)$ el conjunto de los operadores hermitianos de E, observemos primero que $u = u^*$ implica $(\lambda u)^* = \bar{\lambda} u$, luego $\mathcal{H}(E)$ no es un subespacio vectorial de $\mathcal{L}(E)$, es un subespacio vectorial de $\overline{\mathcal{L}(E)}$. Designemos por $\mathcal{H}_2(E;\, \mathbf{C})$ el conjunto de las formas hermitianas sobre E. Naturalmente estos calificativos, operadores hermitianos, formas hermitianas, son relativos a la forma fundamental

$$(x, y) \to f_0(x, y) = (x \mid y)$$

elegida sobre E. A todo elemento u de $\mathcal{H}(E)$ la formula

$$[\forall\, (x, y) \in E^2] \qquad f_u(x, y) = (x \mid u(y))$$

asocia una aplicación de E^2 en $\mathbf{C}$, se ve fácilmente que f_u es una forma hermitiana sobre E, es decir, un elemento de $\mathcal{H}_2(E;\, \mathbf{C})$. Operando como en el § 229, *b*), es decir, utilizando una base (a_i) ortonormal se verá que la aplicación $u \to f_u$ es *biyectiva*; pongamos

$$A = M[u, (a_i)], \qquad X = M[x, (a_i)], \qquad Y = M[y, (a_i)]$$

se tendrá $({}^t\overline{A} = A)$

$$f_u(x, y) = (x \mid u(y)) = {}^t(\overline{AY})X = {}^t\overline{Y}AX$$

luego (§ 236, *d*) A es también la matriz asociada a f_u en la base (a_i).

Es lo mismo por lo tanto estudiar un operador hermitiano, una forma hermitiana o una matriz hermitiana.

c) Diagonalización de un operador hermitiano. Reducción de las formas hermitianas

Sea u un operador hermitiano de E, espacio hermitiano de dimensión n; hay n valores propios que son *a priori* números complejos. Sea λ uno de ellos, x un vector propio no nulo asociado a λ, $(x \mid u(x))$, es real, luego

$$[x \neq 0,\ u(x) = \lambda x] \Rightarrow (x \mid u(x)) = \lambda \,\|x\|^2 \in \mathbf{R}$$

en consecuencia al ser $\|x\|^2$ no nulo, λ es real.

OBSERVACION

Como toda matriz *simétrica real* es un caso particular de matriz *hermitiana,* resulta de lo anterior que los valores propios de una matriz *simétrica real,* por tanto de un operador simétrico de un espacio euclídeo E, son reales. La demostración dada en el § 233 es formalmente idéntica a ésta debido a que se utilizan los vectores x y $\bar{x}$ del complexificado E′ de E.

Por otra parte, al ser x un vector propio no nulo asociado al valor propio λ, se tiene

$$(x \mid y) = 0 \Rightarrow (x \mid u(y)) = (u(x) \mid y) = (\lambda x \mid y) = \lambda(x \mid y) = 0$$

por tanto si y es ortogonal a x, vector propio no nulo de u también lo es $u(y)$.

Finalmente si x_1 y x_2 son dos vectores propios, respectivamente, asociados a λ_1 y λ_2 distintos

$$(u(x_1) \mid x_2) = (x_1 \mid u(x_2))$$

da

$$(\lambda_1 x_1 \mid x_2) = \lambda_1(x_1 \mid x_2) = (x_1 \mid \lambda_2 x_2) = \lambda_2(x_1 \mid x_2)$$

pues λ_1 y λ_2 son reales, de donde

$$[\lambda_1 \neq \lambda_2,\ (\lambda_1 - \lambda_2)(x_1 \mid x_2) = 0] \Rightarrow (x_1 \mid x_2) = 0.$$

Teorema 17′.—*Para todo operador hermitiano u de un espacio hermitiano* E *de dimensión n:*

1. *Los n valores propios de u son reales.*
2. *El subespacio ortogonal a todo vector propio no nulo de u es estable por u.*
3. *Los subespacios propios asociados a los valores propios distintos son ortogonales.*

No siendo isótropo en el espacio hermitiano ningún vector no nulo E, la demostración de la diagonalización de todo operador hermitiano es idéntica a la demostración dada en el § 232, *b*) para la diagonalización de un endomorfismo simétrico de un espacio euclídeo, de donde:

Teorema 18′.—*Siendo u un operador hermitiano de un espacio hermitiano* E *de dimensión n, existen bases ortogonales de* E *formadas por vectores propios de* E.

Al ser unitaria toda matriz de paso de una base ortogonal del espacio hermitiano E a otra base ortogonal de E, tenemos:

Corolario 1.—*Para toda matriz hermitiana* A *de orden n, existe una matriz unitaria* P *tal que* $B = P^{-1}AP$ *es diagonal.*

Finalmente a toda forma hermitiana f se puede asociar de manera biyectiva un operador hermitiano u, tal que

$$[\forall (x, y) \in E^2] \qquad f(x, y) = (x \mid u(y)) = {}^t\bar{Y}AX$$

siendo A hermitiana asociada a f y a u en una misma base ortonormal de E. Si se elige una base ortogonal formada de vectores propios, se tendrá, siendo B ortogonal,

$$f(x, y) = {}^t\overline{Y}BX = \sum_{i=1}^{n} \lambda_i x^i \overline{y^i}, \qquad q(x) = f(x, x) = \sum_{i=1}^{n} \lambda_i \mid x^i \mid^2.$$

La operación precedente se llama *"reducción de las formas hermitianas"* en una base ortonormal de E, relativamente a la forma hermitiana fundamental f_0 de E. La base en la que la matriz de f (o de u) es diagonal es, pues, ortonormal relativamente a f_0 y ortogonal relativamente a f, de donde:

COROLARIO 2. — *Siendo* E *un espacio hermitiano de dimensión n, f_0 una forma hermitiana no degenerada positiva sobre* E *y f una forma hermitiana cualquiera sobre* E, *existen bases ortonormales relativamente a f_0 y ortogonales relativamente a f.*

EJERCICIOS

1. Designemos por $\lambda_1, \ldots, \lambda_m$ los $m(m \leq n)$ valores propios distintos dos a dos de un operador hermitiano del espacio E de dimensión n y por $V(\lambda_1), \ldots, V(\lambda_m)$ los subespacios propios asociados. Demostrar el teorema 18′, demostrando que

$$E = V(\lambda_1) \oplus \ldots \oplus V(\lambda_m).$$

2. Diagonalizar la matriz

$$\begin{pmatrix} 3 & 2+2i \\ 2-2i & 1 \end{pmatrix}.$$

Ejercicios

501. Siendo E un espacio vectorial sobre el cuerpo conmutativo **K** de característica $\neq 2$, demostrar que toda aplicación q de E en **K** verificando:

1. $(\forall \lambda \in \mathbf{K})(\forall x \in E) \quad q(\lambda x) = \lambda^2 q(x)$.
2. $(x, y) \rightarrow \frac{1}{2}[q(x+y) - q(x) - q(y)]$ es una forma bilineal sobre E,

es una forma cuadrática sobre E.

502. Siendo f una forma bilineal simétrica y q la forma cuadrática asociada, demostrar que para todo x y todo y

$$q(x+y) + q(x-y) = 2q(x) + 2q(y)$$
$$q(x+y) - q(x-y) = 4f(x, y).$$

Interpretar geométricamente estas relaciones cuando E es el espacio euclídeo de dimensión 2.

503. Siendo E un espacio vectorial sobre **R**, se considera una aplicación q de E en **R** verificando:

1. $(\forall(x, y) \in E^2) q(x+y) + q(x-y) = 2q(x) + 2q(y)$.
2. $(\forall(x, y) \in E^2)$ la aplicación de **R** en **R** definida por $\lambda \rightarrow q(x + \lambda y)$ es continua.

Demostrar que la aplicación f de E^2 en **R** definida por

$$(x, y) \rightarrow f(x, y) = \frac{1}{4}[q(x+y) - q(x-y)]$$

es una forma bilineal simétrica sobre E. ¿Cuál es la forma cuadrática asociada a f?

504. Se considera una forma cuadrática q sobre E espacio vectorial de dimensión n sobre **K** cuerpo de característica $\neq 2$. Se pone

$$q(x) = \Phi(x^1, x^2, \ldots, x^n) = \sum_{i=1}^{n} \alpha_{ii}(x^i)^2 + 2 \sum_{1 \leq i < j \leq n} \alpha_{ij} x^i x^j.$$

a) Si $\alpha_{11} \neq 0$, demostrar que existe una forma lineal l sobre E y una forma cuadrática q_1 sobre E, verificando $q_1(x) = \psi(x^2, x^3, \ldots, x^n)$, tal que

$$(1) \qquad q(x) = \frac{1}{\alpha_{11}}[l(x)]^2 + q_1(x).$$

Expresar $l(x)$ con la ayuda de $\Phi'_{x1}(x^1, x^2, \ldots, x^n)$. (Considerar Φ como un trinomio de segundo grado en x^1).

b) Si $\alpha_{11} = 0$ y $\alpha_{12} \neq 0$, demostrar que existen dos formas lineales l_1, l_2 sobre E y una forma cuadrática q_1 verificando $q_1(x) = \psi(x^3, x^4, \ldots, x^n)$ tales que

$$(2) \qquad q(x) = \frac{2}{\alpha_{12}} l_1(x) l_2(x) + q_1(x).$$

Expresar $l_1(x)$ y $l_2(x)$ valiéndose de $\Phi'_{x1}(x^1, x^2, \ldots, x^n)$ y $\Phi'_{x2}(x^1, x^2, \ldots, x^n)$. (Escribir $\Phi(x^1, x^2, \ldots, x^n) - 2\alpha_{12}x^1x^2$ separando los términos que contienen x^1 y los términos que contienen x^2).

c) Deducir de las fórmulas (1) y (2) un método de descomposición de una forma cuadrática en cuadrados linealmente independientes llamado método de Gauss. ¿Cuántos cuadrados linealmente independientes se han obtenido? (Si se utiliza la fórmula (2) se observará que

$$4l_1(x)\ l_2(x) = [l_1(x) + l_2(x)]^2 - [l_1(x) - l_2(x)]^2$$

y se demostrará que l_1 y l_2 son independientes así como $l_1 + l_2$ y $l_1 - l_2$).

d) Demostrar que el método precedente es también un método de construcción de una base ortogonal relativa a la forma cuadrática tratada.

En los ejercicios siguientes (505 al 510) *se toma* $\mathbf{K} = \mathbf{R}$. *Se descompondrá las formas cuadráticas dadas sobre* $\mathbf{R}^n (n = 3$ o $4)$ en cuadrados independientes aplicando el método precedente.

505. $(x^1)^2 + (x^2)^2 + (x^3)^2 - 4(x^2x^3 + x^3x^1 + x^1x^2)$.

506. $(x^1)^2 + (x^2)^2 + (x^3)^2 - (x^2x^3 + x^3x^1 + x^1x^2)$.

507. $(x^1)^2 + 6(x^2)^2 - 4x^1x^2 + 8x^3x^1$.

508. $2(x^1)^2 + 3(x^2)^2 - (x^3)^2 - 8x^3x^1$.

509. $$\sum_{1 \leq i < j \leq n} x^i x^j \qquad (n = 3 \quad \text{o} \quad n = 4).$$

510. $(x^1)^2 + (x^2)^2 - 3(x^4)^2 + 5x^1x^2 - 3x^1x^4 + 2x^2x^4 - 7x^3x^4$.

En los ejercicios siguientes (511 al 513) *se hace* $\mathbf{K} = \mathbf{R}$. *Se buscará el rango de las formas cuadráticas dadas discutiendo según el valor de los párrafos* λ *y* μ. *En cada uno de los casos que se hallará se descompondrá en seguida cada forma cuadrática en cuadros independientes.*

511. $(1 - \lambda)(x^1)^2 + 2\mu x^1x^2 + (1 + \lambda)(x^2)^2 - 2\lambda x^1x^3 + 2\mu x^2x^3 + \mu(x^3)^2$.

512. $(x^1)^2 + (x^2)^2 + 2x^3(x^1 \cos \lambda + x^2 \operatorname{sen} \lambda)$.

513. $(x^1)^2 - 3(x^2)^2 - 4(x^3)^2 + \lambda(x^4)^2 + 2\mu x^1x^2$.

514. Siendo f una forma bilineal simétrica no degenerada sobre E de dimensión finita sobre $\mathbf{K}$, demostrar que cualquiera que sean los subespacios vectoriales F y G de E

$$(F + G)^\perp = F^\perp \cap G^\perp, \qquad (F \cap G)^\perp = F^\perp + G^\perp.$$

515. Sea f una forma bilineal simétrica no degenerada sobre E de dimensión finita sobre $\mathbf{K}$. Sea F un subespacio de E no isótropo.

a) Demostrar que existe un endomorfismo de E único p_F tal que para todo x

$$p_F(x) \in F, \qquad x - p_F(x) \in F^\perp$$

p_F se le llama el *operador de proyección ortogonal sobre* F.

b) Demostrar que

$$p_F \circ p_F = p_F, \qquad p_F(E) = F, \qquad p_F(F^\perp) = \{0\}.$$

c) Demostrar que p_F es un operador simétrico relativamente a f. ¿En qué caso puede ser ortogonal relativamente a f? Deducir de lo anterior que no es correcto aplicar el término «proyector ortogonal» para designar el operador de proyección ortogonal.

d) Siendo a un vector no isótropo se pone $F = \mathbf{K}a$ y $G = F^{\perp}$. Demostrar

$$p_F(x) = f(x, a)\,[f(a, a)]^{-1}a,$$
$$p_G(x) = x - p_F(x).$$

516. Siendo f una forma bilineal simétrica no degenerada sobre E de dimensión finita sobre **K** de característica $\neq 2$, se considera un automorfismo u de E tal que $u^2 = e$ y se pone $2v = e - u$, $2w = e + u$. Demostrar que las propiedades siguientes son equivalentes:

a) u es un operador ortogonal relativamente a f.

b) $v(E)$ y $w(E)$ son ortogonales relativamente a f.

c) u es autoadjunto relativamente a f.

(Se considera Im v, Im w, Ker v, Ker w y se utilizarán los ejercicios 157 y 158 del capítulo 7.)

517. Siendo f una forma bilineal simétrica no degenerada sobre E de dimensión finita sobre **K** de característica $\neq 2$ se considera un subespacio F de E, no isótropo.

a) Demostrar que existe un automorfismo de E único S_F tal que

$$s_F(x) = x \quad \text{para} \quad x \in F, \qquad s_F(x) = -x \quad \text{para} \quad x \in F^{\perp}.$$

Demostrar que S_F es ortogonal relativamente a f. s_F se llama simetría respecto a F. (Observar que $(s_F)^2 = e$.)

b) Siendo H un hiperplano no isótropo a E, demostrar que para *todo* x de E

$$s_H(x) = x - 2f(x, a([f(a, a)]^{-1}a)$$

siendo a un vector no nulo ortogonal a H. (Utilizar el ejercicio 515, **d**.)

518*. Siendo f una forma bilineal simétrica no degenerada sobre E de dimensión n sobre **K** de característica $\neq 2$ se propone demostrar que todo automorfismo u ortogonal relativamente a f está compuesto al menos de n simetrías (V. ej. 317) respecto a los hiperplanos no isótropos de E (es decir, que el grupo $\mathbf{O}(f)$ está engendrado por simetrías). Se razonará por recurrencia sobre n. Siendo x un vector no isótropo de E se estudiará sucesivamente los casos siguientes:

a) $u(x) = x$. Si H es el hiperplano ortogonal a x, se demostrará que $u(H) = H$ y se considerará la aplicación u' inducida por u sobre H.

b) $u(x) = -x$. Se reducira al caso anterior escribiendo $v = s \circ u$, siendo s la simetría respecto a H ortogonal a x.

c) Caso general: $y = u(x)$. Se demostrará que $a = x - y$ o $b = x + y$ es no isótropo y se reducirá al caso **a** (o **b**) considerando la simetría s respecto al hiperplano H ortogonal a a (o a b).

519. Siendo f una forma bilineal simétrica sobre E de dimensión finita se dice que F subespacio de E es *totalmente isótropo* relativamente a f si para todo x de F, $f(x, x) = 0$.

a) Demostrar que para que F sea totalmente isótropo es necesario y suficiente siendo a un vector no nulo ortogonal a H. (Utilizar el ejercicio 515, **b**.)

b) Demostrar que todo subespacio de F totalmente isótropo es totalmente isótropo.

c) Siendo F y G totalmente isótropos. ¿Qué serán $F \cap G$ y $F + G$? (estudiar en particular el caso en que F y G son ortogonales).

d) Si F y G son solamente isótropos, ¿qué se puede decir de $F \cap G$ o de $F + G$? (tomar ejemplos en $E = C^3$ provisto de $q(x) = (x^1)^2 + (x^2)^2 + (x^3)^2$).

520*. Tomamos las notaciones del § 229, *b*), se propone demostrar *sin recurrir a una base ortonormal* de E (que puede no existir) que $S_f(E)$, espacio vectorial de los endomorfismos de E, simétricos relativamente a f, es isomorfo a $S_2(E; \mathbf{K})$ espacio vectorial de las formas bilineales simétricas sobre E.
Se designará por Φ la aplicación de $S_f(E)$ en $S_2(E; \mathbf{K})$ definida por

$$u \to \Phi(u) = f_u, \qquad f_u(x, y) = [u(x), y].$$

a) Demostrar que Φ es lineal.

b) Demostrar que Φ es inyectiva.

c) Demostrar que Φ es suprayectiva. (Si $g \in S_2(E; \mathbf{K})$ se pondrá $g_x = g(x, \ .\ ) \in E^*$ y se introducirá x' único tal que para todo y de E, $g_x(y) = f_{x'}(y)$ (V. § 224). Habiendo puesto $x' = \nu(x)$, se demostrará que $\nu \in S_f(E)$).

d) Demostrar que u y f_u tienen el mismo rango.

521. Determinar las signaturas de las formas cuadráticas dadas en los ejercicios 505 al 510.

522*. Sea E un espacio vectorial de dimensión n sobre $\mathbf{R}$ provisto de una forma bilineal simétrica no degenerada positiva f. Se considera un *proyector autoadjunto* de E, es decir, un endomorfismo p de E tal que $p = p^2 = p^*$.

a) Demostrar que existe un subespacio F de E tal que p sea el operador de proyección ortogonal p_F (V. ej. 515).

b) F y G son dos subespacios de E se designa por F′ (resp. G′) el subespacio de F (resp. G) descrito por los vectores ortogonales a $F \cap G$.
Demostrar que p_F y p_G permutan si y sólo si F′ y G′ son ortogonales y que en este caso

$$p_{F \cap G} = p_F \circ p_G.$$

Hallar también un resultado clásico cuando F y G son dos planos del espacio euclídeo $\mathbf{R}^3$.

523. Sea E un espacio euclídeo y F y G dos hiperplanos de E. Se consideran las simetrías s_F y s_G (V. ej. 517). ¿En qué condición permutan s_F y s_G?
¿A qué es igual entonces $s_F \circ s_G$? (utilizar ej. 515, **d**) y 522).

524*. Sea E el espacio vectorial sobre $\mathbf{R}$ de las funciones reales definidas y continuas sobre $[a, b]$. Se considera la aplicación de φ de E^2 en $\mathbf{R}$ definida por

$$\varphi(f, g) = \int_a^b f(x)g(x)\,dx.$$

a) Demostrar que φ es una forma bilineal simétrica no degenerada positiva sobre E. Se dirá que una sucesión (f_i) es ortogonal si para $i \neq j$, $\varphi(f_i, f_j) = 0$ y ortonormal si $\varphi(f_i, f_i) = 1$ para todo i.

b) Demostrar que $d(f, g) = \| f - g \| \geq 0$ definida por

$$\| f - g \|^2 = \varphi(f - g, f - g)$$

es una distancia sobre E.

c) Sea (g_i) una sucesión finita ortonormal, se llama G_n el subespacio vectorial de E engendrado por $\{g_0, g_1, \ldots, g_n\}$. ¿Cuál es la dimensión de G_n? Para todo $i \geq 0$

se pone $\alpha_i = \varphi(f, g_i)$, siendo f un elemento cualquiera de E. Sea $g = \lambda_0 g_0 + \ldots + \lambda_n g_n$ un elemento cualquiera de G_n, demostrar que

$$\| f - g \|^2 = \| f \|^2 - (\alpha_0^2 + \ldots + \alpha_n^2) + (\lambda_0 - \alpha_0)^2 + \ldots + (\lambda_n - \alpha_n)^2.$$

Deducir los valores que hay que dar a $\lambda_0, \ldots, \lambda_n$ para que $\| f - g \|$ tenga el menor valor posible. Demostrar que $\alpha_0^2 + \ldots + \alpha_n^2 \leq \| f \|^2$; ¿se puede tener la igualdad?

525*. Tomamos las mismas notaciones que en el ejercicio anterior con $[a, b] = [-1, 1]$ se propone determinar una sucesión ortogonal de polinomios (P_n) tal que P_n sea de grado n y que $P_n = 1$ para todo $n \geq 0$.

a) Calcular P_0, P_1, P_2 y demostrar de una manera general que se puede calcular P_n de modo único.

Demostrar que P_n es ortogonal a todo monomio $x^m (m < n)$. De lo anterior deducir que P_n es ortogonal a todo polinomio **R** de grado $< n$.

b) Para n fijo, demostrar que existe un polinomio Q y uno sólo de grado $2n$, divisible por $(x - 1)^n$ y tal que $P_n = \dfrac{d^n Q}{dx^n}$. Escribiendo que P_n es ortogonal sucesivamente a cada monomio $1, x, \ldots, x^{n-1}$, demostrar que Q es divisible por $(x + 1)^n$. (Se calculará por partes $\varphi[x^m, Q^n(x)]$ para $m = 0, 1, \ldots, n - 1$, utilizando el hecho de que 1 es raíz de orden n de Q.)

Deducir que:

$$Q(x) = k_n(x^2 - 1)^n.$$

Calcular la constante k_n. (M.G.P.)

526. Siendo E el espacio vectorial sobre **R** de los polinomios con coeficientes reales y f una aplicación real, continua, estrictamente positiva sobre $[-1, +1]$ se considera la aplicación F de E^2 en **R** definida por

$$F(P, Q) = \int_{-1}^{+1} P(x)Q(x)f(x)\, dx.$$

Se designa por E_k el subespacio de E descrito por los polinomios de grado *estrictamente inferior* a k.

a) Demostrar que F es una forma bilineal simétrica no degenerada, positiva sobre E. Se dirá que P y Q son ortogonales y solamente si $F(P, Q) = 0$.

b) Sea $P_0 = 1, P_1, \ldots, P_n, \ldots$ una sucesión de polinomios de E, unitarios y de grado igual a su índice, demostrar que las tres condiciones siguientes definen una sola y una misma sucesión de polinomios (P_n):

1. P_h y P_k son ortogonales cualesquiera que sean $h \neq k$ (estando $P_0, \ldots, P_{k-1}$ determinados, observar que $P_k(x) - x^k$ pertenece a E_k).
2. Cualquiera que sea k, P_k es ortogonal a todo polinomio A de grado $< k$.
3. B_k es un polinomio unitario de grado k, $F(B_k, B_k)$ es mínimo para $B_k = P_k$ cualquiera que sea k. (Se observará que $B_k - P_k$ pertenece a E_k.)

c) Se designa por **R** el polinomio definido por

$$R(x) = (x - x_1)(x - x_2) \ldots (x - x_h),$$

donde $x_1, \ldots, x_h$ son los ceros de P_k distintos 2 a 2 reales estrictamente comprendidos entre -1 y $+1$ y de orden impar. Demostrar que $F(P_k, R) \neq 0$. Deducir el grado

de R y que P_k tiene todas sus raíces reales distintas y comprendidas estrictamente entre -1 y $+1$.

d) Demostrar que existen números reales a_n y b_n tales que

$$P_n(x) = (x + a_n)P_{n-1}(x) + b_nP_{n-2}(x).$$

(Se observará que $P_n(x) - xP_{n-1}$ pertenece a E_n y se expresará a_n y b_n con la ayuda de integrales definidas.)

Demostrar que si la función f es par $a_n = 0$ y que la función polinomio P_n es par o impar según que n sea par o impar.

e) Si Q_n es una sucesión de polinomios de grado igual a su índice tal que Q_h y Q_k sean ortonormales cualesquiera que sean $h \neq k$. Demostrar que $Q_n = \lambda_n P_n$, donde λ_n es un número real no nulo arbitrario.

En los ejercicios siguientes del 527 al 531, A es la matriz de un endomorfismo simétrico u de un espacio euclídeo E referido a una base ortonormal (a_i). Determinar una base ortonormal (b_i) teniendo la misma orientación que (a_i) y tal que $B = M(f, (b_i))$ sea diagonal. Determinar la matriz de paso P de (a_i) a (b_i). Escribir en la base (a_i) y la base (b_i) la expresión de $q(x)$, donde q es la forma cuadrática asociada a u.

527.

$$A = \begin{pmatrix} 4 & 1 & 1 \\ 1 & 4 & 1 \\ 1 & 1 & 4 \end{pmatrix}.$$

528.

$$A = \begin{pmatrix} -2 & -2 & 1 \\ -2 & 1 & -2 \\ 1 & -2 & -2 \end{pmatrix}$$

529.

$$A = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}.$$

530.

$$A = \begin{pmatrix} a & b & b \\ b & a & b \\ b & b & a \end{pmatrix}, \qquad A = \begin{pmatrix} a & b & b & b \\ b & a & b & b \\ b & b & a & b \\ b & b & b & a \end{pmatrix} \qquad (a, b \in \mathbf{R}).$$

531*. Generalizar el ejercicio anterior a $A = (\alpha_i^j)$ de orden n con $\alpha_i^i = a$ para todo i y $\alpha_i^j = b$ para todo $i \neq j$.

En los ejercicios siguientes del 532 al 536, q es una forma cuadrática definida sobre un espacio euclídeo E por el valor $q(x)$ en función de las coordenadas x^i de x respecto a una base (a_i) ortonormal de E. Hallar la forma reducida de cada una de estas formas cuadráticas en una base ortonormal (b_i) de E; indicar cada vez la matriz de paso de (a_i) a (b_i).

532. $(x^1)^2 + (x^2)^2 - 5(x^3)^2 + 2x^2x^3 - 2x^3x^1 - 6x^1x^2$.

533. $(x^1)^2 - 2(x^2)^2 - (x^3)^2 + 2x^1x^2$.

534. $2(x^1)^2 + 2(x^2)^2 + (x^3)^2 - 2x^3x^1 - 2x^1x^2$.

535. $(x^1)^2 + (x^2)^2 + 2x^3(x^1 \cos\alpha + x^2 \operatorname{sen}\alpha)$.

536 $2x^1x^2 - 6x^1x^2 - 6x^2x^4 + 2x^3x^4$.

En los siguientes ejercicios del 537 al 540 demostrar que A es una matriz que representa una rotación del espacio euclídeo de dimensión tres, referida a una base ortonormal. Determinar el eje y el ángulo de cada una de estas rotaciones.

537.

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

538.

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

539.

$$A = \frac{1}{9}\begin{pmatrix} 8 & 1 & -4 \\ -4 & 4 & -7 \\ 1 & 8 & 4 \end{pmatrix}.$$

(École Polytechnique, oral).

540.

$$A = \begin{pmatrix} b^2 + a^2 \cos t & a \operatorname{sen} t & ab(1 - \cos t) \\ -a \operatorname{sen} t & \cos t & b \operatorname{sen} t \\ ab(1 - \cos t) & -b \operatorname{sen} t & a^2 + b^2 \cos t \end{pmatrix}$$

donde a, b, t son números reales tales que $a^2 + b^2 = 1$.

(Concurso de Minas, extractado)

541. Demostrar que toda matriz ortogonal real de orden 2 pertenece a uno de los tipos ($\theta \in \mathbf{R}$)

$$A_1 = \begin{pmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{pmatrix} \qquad A_2 = \begin{pmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{pmatrix}$$

Determinar para cada caso los valores propios y los vectores propios. Interpretar geométricamente los operadores definidos por estas matrices relativamente a una base ortonormal del espacio vectorial euclídeo de dimensión 2.

542*. Sea u operador ortogonal del espacio vectorial euclídeo de dimensión n.

a) Demostrar que si x es un vector propio no nulo de u, el hiperplano ortogonal a u es invariante por u y más generalmente que si F es un subespacio de E invariante por u, lo mismo ocurre con $F^{\perp}$.

Demostrar que dos vectores propios asociados a dos valores propios distintos son ortogonales.

b) ¿Cuáles son los valores propios reales de u? (V. § 228, ej. 3). Demostrar que si todos los valores propios de u son reales, existe una base ortonormal de E tal que en esta base la matriz de u sea

$$\begin{pmatrix} I_p & 0 \\ 0 & -I_p \end{pmatrix} \qquad (p \geq 0,\ q \geq 0,\ p + q = n)$$

(Se demostrará que $E = V(1) \oplus V(-1)$.)

c) ¿Cuáles son los valores propios no reales de u (V. § 232, ej. 2). Demostrar que si ningún valor propio de u es real $n = 2r$ y que existe una base ortonormal de E tal que la matriz de u, en esta base, sea un cuadro diagonal de r matrices de orden 2 de la forma

$$A_h = \begin{pmatrix} \cos\theta_h & -\operatorname{sen}\theta_h \\ \operatorname{sen}\theta_h & \cos\theta_h \end{pmatrix}$$

$(\theta_h)(1 \leq h \leq r)$ es una familia de números reales no múltiplos de π. (Se demostrará que $E = R_1 \oplus \ldots \oplus R_r$, siendo $R_h (1 \leq h \leq r)$ de dimensión 2, invariante por u así como $(R_h)^{\perp}$ y se utilizará el ejercicio 541 después de haber considerado al operador u_h inducido por u sobre R_h).

d) *Demostrar en el caso general que existe una base ortonormal de* E *tal que en esta base la matriz de* u *sea*

$$\begin{pmatrix} I_p & & & & 0 \\ & -I_q & & & \\ & & A_1 & & \\ & & & \ddots & \\ 0 & & & & A_r \end{pmatrix} \qquad \begin{matrix} (p \geq 0 \quad q \geq 0 \quad r \geq 0) \\ (p + q + 2r = n) \end{matrix}$$

donde (A_h) es una familia de matrices del tipo definido en **c**).

543. *Demostrar que un endomorfismo* u *de* E_n *espacio euclídeo orientado de dimensión* n es una rotación si existen dos base ortonormales de la misma orientación (a_i) y (b_i) tales que $u(a_i) = b_i$ $(1 \leq i \leq n)$.

¿Cuál es el número máximo p_n de parámetros reales de los que depende una rotación de E_n? (Razonar por recurrencia.) Dar los valores de p_2, p_3, p_4.

544. Si (a, b, c) y (a', b', c') son dos bases de igual orientación del espacio euclídeo orientado E_3 se considera la rotación r definida por $r(a) = a'$, $r(b) = b'$, $r(c) = c'$. Demostrar que hay tres rotaciones r_1, r_2, r_3 tales que $r = r_3 \circ r_2 \circ r_1$ con

r_1 rotación alrededor de c del ángulo φ que envía (a, b, c) en (a_1, b_1, c_1)

r_2 rotación alrededor de a_1 del ángulo θ que envía (a_1, b_1, c_1) en (a_2, b_2, c_2)

r_3 rotación alrededor de c_2 del ángulo ψ que envía (a_2, b_2, c_2) en (a', b', c').

Calcular la matriz r de respecto a la base (a, b, c) mediante los números φ, θ, γ llamados ángulos de EULER de la rotación r, o de la matriz $M(r)$.

(Tomar a_1 en la intersección de los planos definidos, respectivamente, por a, b y a', b'.)

545. Tenemos dos vectores v_1 y $v_2 (v_1 \neq 0)$ del espacio euclídeo orientado E_3 de dimensión 3. Demostrar que el *producto vectorial* $w = v_1 \wedge v_2$ puede obtenerse efectuando sucesivamente las tres operaciones siguientes:

1. Proyección ortogonal de v_2 sobre el plano (pasando por 0 ortogonal a a_1). Sea v_2' el vector obtenido.
2. Rotación de v_2' alrededor de v_1, de ángulo $+\pi/2$. Sea v_2'' el vector obtenido.
3. $w = \| v_1 \| v_2''$.

546. En E_3, espacio euclídeo orientado de dimensión 3, se designa por v' el vector imagen de un vector v en la rotación r del ángulo θ alrededor de un vector unitario k.

a) Demostrar que si $k \cdot v = 0$ se tiene

(1) $$v' = v \cos\theta + (k \wedge v) \operatorname{sen}\theta$$

(utilizar el ejercicio precedente).

b) Sean v_1 y v_2 las proyecciones ortogonales respectivas de v sobre k y sobre el plano (pasando por 0) ortogonal a k. Demostrar que

$$v_1 = (k \cdot v)k, \qquad v_2 = v - (k \cdot v)k.$$

c) Demostrar en el caso general la fórmula

(2) $$v' = v \cos\theta + (k \wedge v) \operatorname{sen}\theta + (1 - \cos\theta)(k \cdot v)k.$$

d) Sean α, β, γ las coordenadas de k en una base ortonormal de sentido directo (a_1, a_2, a_3) de E_3. Demostrar que en esta base la matriz de r es

$$\begin{pmatrix} \cos\theta + (1-\cos\theta)\alpha^2 & -\gamma \operatorname{sen}\theta + (1-\cos\theta)\alpha\beta & \beta \operatorname{sen}\theta + (1-\cos\theta)\alpha\gamma \\ \gamma \operatorname{sen}\theta + (1-\cos\theta)\alpha\beta & \cos\theta + (1-\cos\theta)\beta^2 & \alpha \operatorname{sen}\theta + (1-\cos\theta)\beta\gamma \\ -\beta \operatorname{sen}\theta + (1-\cos\theta)\alpha\gamma & \alpha \operatorname{sen}\theta + (1-\cos\theta)\beta\gamma & \cos\theta + (1-\cos\theta)\gamma^2 \end{pmatrix};$$

Calcular la matriz de la rotación correspondiente a

$$k = \frac{a_1 + a_2 + a_3}{\sqrt{3}}, \qquad \theta = \frac{2\pi}{3}.$$

NOTA: Se obtiene así la matriz de la rotación más general en función de los parámetros reales α, β, γ, θ ligados por la relación $\alpha^2 + \beta^2 + \gamma^2 = 1$; esta matriz es, pues, menos simple que la matriz obtenida en el ejercicio 544 en función de los tres ángulos de EULER φ, θ, ψ de la rotación, pero esta matriz tiene la ventaja de poner en evidencia el eje y el ángulo de la rotación.

547. H_1 y H_2 son dos matrices hermitianas del mismo orden, demostrar que las propiedades siguientes son equivalentes:

1. H_1 y H_2 permutan;
2. H_1H_2 es hermitiano;
3. Existe una misma matriz unitaria U tal que $U^{-1}H_1U$ y $U^{-1}H_2U$ sean diagonales (utilizar capítulo 14, ej. 479).

548. A es la matriz asociada a una forma hermitiana f sobre E vectorial de dimensión n sobre **C** en una base (a_i).

a) Demostrar que

$$f(x, y) = Y^*AX$$

X e Y son respectivamente las matrices columnas asociadas a x e y en la base (a_i). (Se recuerda que para toda matriz rectangular compleja M, $M^* = {}^t\bar{M}$.)

b) Demostrar que existe una matriz cuadrada compleja inversible tal que P^*AP sea diagonal (utilizar el hecho que existe una base ortogonal relativamente a f y la observación 1 del § 226).

549. Siendo f una forma hermitiana no degenerada definida sobre E, espacio vectorial de dimensión n sobre **C**, se considera un endomorfismo u de E y siendo (a_i) una base de E se pone

$$A = (f(a_i, a_j)), \qquad M = f(u, (a_i)), \qquad M' = f(u^*, (a_i))$$

demostrar

$$M' = A^{-1}M^*A.$$

(Utilizar el ejercicio anterior.)

550. Se dice que una matriz hermitiana compleja (resp. simétrica real) de orden n es positiva si en una base de E, espacio vectorial de dimensión n sobre **C** (resp. sobre **R**) está asociada a una forma hermitiana (resp. bilineal simétrica) positiva.

a) Si A es una matriz cuadrada compleja demostrar que A^*A es una matriz hermitiana positiva. ¿Qué se puede decir de los valores propios de A^*A?

b) Recíprocamente demostrar que H matriz hermitiana es positiva si existe una matriz compleja A tal que $H = A^*A$ (utilizar ej. 548, **b**).

c) Tratar las preguntas anteriores para S simétrica real positiva.

551. Tenemos la matriz

$$H = \begin{pmatrix} 2 & 1-i \\ 1+i & 3 \end{pmatrix}.$$

Demostrar que existe una sola matriz hermitiana positiva (V. ej. 550) H' tal que $H = H'^2$. Calcular H'.

552*. Siendo H una matriz hermitiana positiva (V. ej. 550) demostrar que existe una matriz hermitiana positiva única H' tal que $H = H'^2$. (Observar que H y H' permutan y utilizar el ejercicio 547.)

553. Sea A una matriz de $\mathbf{GL}_n(\mathbf{C})$; se considera la matriz hermitiana A^*A positiva (ver el ejercicio 550).

a) Demostrar que existe una matriz unitaria U y una matriz hermitiana H positiva con determinante no nulo tales que $A = UH$ (se introducirá la matriz hermitiana positiva B tal que $A^*A = B^2$, ver ejercicio precedente).

b) Calcular U y H para $A = \begin{pmatrix} 2 & i \\ 2+3i & 1-i \end{pmatrix}$.

(Comparar este ejercicio con el ejercicio 463, capítulo 14.)

554*. Se dice que un operador u de un espacio hermitiano E de dimensión finita sobre **C** es *normal* si $u \circ u^* = u^* \circ u$.

a) Demostrar que u es normal si y solamente si $u - \alpha e$ es normal ($\alpha \in \mathbf{C}$, e identidad de E).

b) Demostrar que $\operatorname{Ker} u = \operatorname{Ker} u^*$ si u es normal.

c) Demostrar que si λ es valor propio de u normal, $\bar{\lambda}$ es valor propio de u^*.

d) Demostrar que los subespacios propios $V(\lambda)$ y $V(\lambda')(\lambda \neq \lambda')$ de u normal son ortogonales.

e) Demostrar que todo subespacio propio de u normal sea $V(\lambda)$ es estable por u^* y que el ortogonal de $V(\lambda)$ en el espacio hermitiano E es estable por u y u^*.

f) Deducir de los resultados anteriores que para todo operador normal de E existe una base ortogonal de vectores propios. (Observar que este resultado se aplica a los operadores hermitianos como lo hemos demostrado en el § 240, **c)** y también a los operadores unitarios.)

555*. Se dice que una matriz compleja A es *normal* si $A^*A = AA^*$. Sean dos matrices complejas A y B normales tales que $AB = BA$, demostrar que existe una matriz unitaria U tal que $U^{-1}AU$ y $U^{-1}BU$ sean diagonales (utilizar cap. 14, ej. 479 y el ejercicio precedente).

556. Sea H una matriz hermitiana compleja de orden n.

a) Demostrar que $H - iI_n$ es inversible y que

(1) $$U = (H + iI_n)(H - iI_n)^{-1}$$

es unitario. Calcular los valores propios de U en función de los valores propios de H. Comprobar que son de módulo 1, pero que la matriz unitaria U dada por (1) está desprovista del valor propio 1.

b) Siendo U una matriz unitaria desprovista del valor propio 1 demostrar que

(2) $$H = i(U + I_n)(U - I_n)^{-1}$$

es hermitiana.

Las fórmulas (1) y (2) de CAYLEY demuestran la relación que existe entre matrices unitarias y matrices hermitianas.

557. Se dice que una matriz A de $M_n(\mathbf{C})$ es antihermitiana si $A^* = -A$.

a) Demostrar que A es antihermitiana si y sólo si iA es hermitiana.

b) Demostrar que toda matriz M de $M_n(\mathbf{C})$ se escribe de una manera única de la forma

$$M = H + A$$

o bien en la forma

$$M = H + iH'$$

siendo A antihermitiana y H y H′ hermitianas.

c) Demostrar que E es normal (V. ej. 555) si y sólo si las matrices H y H′ permutan.

558. a) Demostrar que todas las matrices del grupo $\mathbf{SU}(2, \mathbf{C})$ son de la forma

$$A = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}.$$

donde a y b son dos números complejos ligados por la relación $a\bar{a} + b\bar{b} = 1$.

b) Deducir de la pregunta a), que se puede escribir igualmente

$$A = \begin{pmatrix} e^{i\alpha} \cos \varphi & e^{i\beta} \operatorname{sen} \varphi \\ -e^{-i\beta} \operatorname{sen} \varphi & e^{-i\alpha} \cos \varphi \end{pmatrix}$$

en función de tres parámetros reales α, β, φ.

c) Demostrar que los valores propios de A son $e^{i\theta}$, $e^{-i\theta}$ con $\cos \theta = \cos \alpha \cos \varphi$. Determinar los vectores propios.

559. Se consideran los grupos $\mathbf{GL}(n, \mathbf{C})$, $\mathbf{GLS}(n, \mathbf{C})$, $\mathbf{U}(n, \mathbf{C})$, $\mathbf{SU}(n, \mathbf{C})$, $\mathbf{GL}(n, \mathbf{R})$, $\mathbf{GLS}(n, \mathbf{R})$, $\mathbf{O}(n, \mathbf{R})$, $\mathbf{SO}(n, \mathbf{R})$, $\mathfrak{S}_n$ (grupo simétrico), $\mathfrak{A}_n$ (grupo alternado). Colocar estos diferentes grupos en un diagrama en el que el diagrama elemental $G \rightarrow H$ indique que H es un subgrupo de G. ($\mathbf{GLS}(n, \mathbf{K})$ designa el subgrupo de $\mathbf{GL}(n, \mathbf{K})$ descrito por los operadores de determinante $+1$.)

INDICE DE SIMBOLOS

Las cifras remiten a las páginas donde se definen los símbolos

INDICE TERMINOLOGICO

Los números indican las páginas. Cuando un término está definido por un grupo de palabras, este término ha sido clasificado por su primera palabra, excepto los que contienen un nombre propio, que se clasifican por éste.